

# Mediating Conflicts in Minecraft: Empowering Learning in Online Multiplayer Games

Petr Slovak, Katie Salen, Stephanie Ta, Geraldine Fitzpatrick

Human Computer Interaction Group, TU Wien (Vienna University of Technology), Austria  
University of California at Irvine, US

## ABSTRACT

Multiplayer online games, such as Minecraft, have the potential to be powerful sites for youth learning, but can be plagued by inter-personal conflicts. This brings the need for online moderation. However, only very little is known about the practices through which such moderation happens, or how socio-technical systems could be designed to enable ‘safe’ learning spaces online. To start addressing this gap, our research examines the existing mediation practices within a moderated Minecraft server for children aged 8-13. As part of our 14 months long engagement, we triangulate data from participant observation, interviews, and analysis of server logs. We demonstrate how—in trying to ‘keep peace’—the online moderators monopolised the conflict resolution process, essentially preventing the children from actively working with and learning from the experiences of conflict. In response to these findings, we present an alternative framework for online conflict mediation, suggesting ways in which existing conflict resolution techniques originating in Prevention Science could be re-interpreted for online multiplayer settings.

## AUTHOR KEYWORDS

Learning, Conflict Resolution, Multiplayer Online Games, Tweens, Learning Sciences, Social Emotional Learning

## ACM CLASSIFICATION KEYWORDS

H.5.m. [Information interfaces and presentation]: Miscellaneous.

## INTRODUCTION

Multiplayer videogames have become an integral part of children and adolescents’ lives in the Western world: 97% report playing for at least one hour per day in the United States [27, 40]; and the majority of gamers across all age ranges play multiplayer games at least weekly [1]. Research in the social sciences shows that—under the right conditions—gaming with others can have positive social benefits: players can acquire prosocial skills if they are playing games that reward cooperation [26, 63]; and are more likely to be engaged in civic movements in their everyday life if similar aspects are part of the gaming experience [27, 40]. Multiplayer

gaming is however also associated with risks of unmanaged conflicts, such as being the target of anti-social behaviour or ‘griefing’ online (e.g., [5, 23]). These experiences increase frustration [47], can reinforce negative gender stereotypes [5], and limit the learning potential of the online spaces [36].

Encountering such ‘toxicity’ online is particularly troubling when young gamers are involved [11, 27, 37]. The perception of such risks leads parents and educators to have concerns about their children’s online safety, including who they are interacting with and what they might experience. This has resulted in calls for ‘moderated’ online spaces, which have technical and social controls in place to safeguard young gamers (cf., [36, 46]). However, little research has focused on understanding the effects of existing moderating approaches on multiplayer servers, and especially how these practices can support or hinder the potential for learning social skills.

This paper aims to start addressing this gap by outlining how the commonly used deterrence-based approaches to moderation on a multiplayer server can disempower (rather than support) young players in developing social skills online, and to suggest an alternative, learning-driven framework for online moderation. We develop this line of argument in two steps:

First, we report on a series of research engagements within the context of a moderated Minecraft server for children aged 8-13 over a period of 14 months. We show how the existing moderation culture in our case-study setting—as well as many other servers—appears to be predominantly grounded in a punitive authoritarian paradigm (surveillance, detection, punishment model) rather than in an empowering paradigm where children are actively supported to work with and learn from the experiences of conflict that naturally arise in an online setting. These arguments draw on a thematic analysis of more than 350 hours of server logs complemented by interviews and long-term participatory engagement with experienced moderators on the server.

Second, we build on existing literature in Prevention Science (cf., [18, 19, 22, 28, 29, 65]) to develop a conceptual grounding for an alternative, ‘competency building’ approach to moderation in online settings. We articulate a set of underlying principles consistent across the reviewed programs and highlight the associated shifts in moderating culture these would imply for the online moderating practices we observed.

Overall, this paper complements the existing toxicity literature (to be discussed below) by suggesting a shift from the predominant authoritarian policing approach towards understanding how moderated online contexts could be designed to serve as powerful sites for transferable learning around social skills and digital citizenship. To do so, the presented results combine (i)

Permission to make digital or hard copies of all or part 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 components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

CHI 2017, May 06 - 11, 2017, Denver, CO, USA

© 2017 Copyright held by the owner/author(s). Publication rights licensed to ACM. ISBN 978-1-4503-4655-9/17/05...\$15.00

DOI: <http://dx.doi.org/10.1145/3025453.3025516>

in-depth ethnographic engagement with a particular community; and (ii) a conceptual framework, drawing on scholarship across learning sciences and prevention science, that can guide future research and design in this space.

## RELATED WORK

While very little work is available on impacts of online moderation on social competencies of players (cf., [46]), the broader set of literatures relevant to conflicts online, multiplayer games, and learning is vast. In what follows, we first introduce the existing work in HCI around toxicity in online multiplayer games and the commonly used approaches to moderation. We then move to present a selective review of scholarship across learning sciences, with a specific focus on games studies and connected learning. Drawing on the existing accounts of how social learning can arise as part of ‘enculturation’ through engagements within a gaming community over time (cf., [25, 32, 43, 48, 57, 60]), it provides an analytical lens that we will use to unpack the moderation strategies observed in our case-study setting.

### Multiplayer games & conflict in HCI

Interplayer conflict plays an integral role in the perception of online gaming spaces as potentially ‘toxic’ environments. Already at CHI’94, an expert panel has pointed to “*the unfortunate fact of life that where there are multi-user computer systems, there will be antisocial behavior*” [17]. The negative impacts of such a ‘culture of toxicity’ online are well-recognised by both academics and commercial organisations alike [4, 23, 38, 56]. The situation is particularly salient within Massive Multiplayer Online (MMO) games, where social ties between players are not stable—such as random matching of players in League of Legends [42]. Antisocial behavior perpetrated by others is the least favorite aspect of online game play as reported by players [30] and can lead to lasting effects: for example, data suggests that League of Legends players who experience in-game toxicity are up to 320% more likely to quit playing the game [41, 50].

In contrast to common expectations that most of toxicity online comes from a small number of ‘trolls’, the literature suggests that, at least within gaming environments, it is a systemic problem that is unlikely to be solved by selective removal of ‘toxic’ members: To stay with the League of Legends example, only about 1% of players were consistently toxic, producing only about 5% of toxicity in game [42]. Still, a uniting feature of the existing strategies to reduce toxic behaviour is the reliance on technological deterrence approaches [10, 38, 42, 49]: the threat of bans or other in-game punishments for infractions, which can be based on crowdsourced decisions by other players (e.g., the ‘Tribunal’ in League of Legends [42]) or machine learning approaches [12, 38].

While such a punitive approach is possibly effective in MMOs like League of Legends with very loose social structure, a similar culture of moderation is also surprisingly present in games and open online worlds specifically designed for teens—such as WebKinz, Neopets, Penguins Club, Whyville and others—where stable social bounds and on-going relationships among peers are among the key features [14, 31, 32, 36, 48]. Social interaction might be either highly restricted to prevent conflicts

altogether (e.g., communication by exchanging predetermined questions and answers in WebKinz [11]), or the community norms expect that players immediately report problems to an adult authority in addition to providing them with technological tools to ‘escape’ any conflict (such as ‘911 calls’ or ‘muting’ and ‘vaporising’ an offending player in Whyville) [37]. Work by Ringland et. al. [46] uncovers a similar dynamic of strict rules and technological restrictions has played out also in more specialised servers, in this case a Minecraft community dedicated for children with autism.

Such reliance on technological, deterrence based ‘solutions’ for interpersonal problems echoes Amy Bruckman’s observation from more than 15 years ago: “*in computer-based communities, it is tempting to throw technological solutions at societal problems [...] technological interventions are [however] rarely more than a band-aid for social problems*” [17]. The learning sciences literature discussed below further supports and extends this critique: the main argument will be that such a punitive culture might not only be a matter of an ‘inefficient band-aid’, but in fact could lead to reinforcing a disempowering set of values and behaviours around social interaction more broadly.

### Games as powerful sites for (social) learning

Learning scientists have studied the relationship between games and learning for decades. Learning in games has been shown to be the result of interaction with an interesting problem context where learners construct meaning [25, 57, 58]. A large body of literature has drawn on the power of games to engage learners with complex concepts such as math, scientific reasoning, or literacy through simulation: offering learners virtual worlds where “their behaviour matters” and mastering a particular curricular concept becomes a problem solving activity within the game-world. These include successful large scale stand-alone games such as Quest Atlantis [7, 8], but also add-on curricula onto commercial games like Minecraft to teach history or programming [24].

What is more pertinent for the questions addressed in this paper is the development of social skills that can appear as a ‘side effect’ of playing multiplayer commercial games [32, 59]: with observations of such phenomena documented across a wide range of titles such as StarCraft [43], Sackboy Planet 2 [45], World of Warcraft [58, 61], and Everquest 2 [62]. Such learning processes are linked to apprenticeship within the *communities of practice* [39] that emerge as part of the social interactions around the game – such as when new players are being ‘shown the ropes’ by other guild members. Through such cognitive apprenticeship [66], the players are ‘enculturated’ into not simply the technical practices of the game but also the dominant cultural perspectives as well: “*this includes adopting the ‘right’ set of values and attitudes toward the game, its content, its goals, world, and other players*” [60, p.123].

Previous work has examined how these processes might impact gamers’ perspectives towards collaborative problem solving, digital media practices, computational literacy as well as existing norms around ‘appropriate’ social interactions in online spaces [32, 59, 61]. However, little is known about the impacts and practices of server-wide moderation on perspectives such as interpersonal interaction and social skills

(cf., [17,36]). The game environment we use to explore these moderation issues is multiplayer Minecraft.

## SETTING

### Multiplayer Minecraft: basics of gameplay, challenges

Minecraft is a sandbox-style video game primarily about creativity and building. Players can modify the terrain of the 3D Minecraft world in many different ways. Because it provides tools for designing, creating, problem-solving, and collaborating, the game also offers significant potential for a range of valuable learning opportunities for players. The game can be played in a number of different modes, the two most popular being Survival and Creative. In Survival mode players try to stay alive by mining and managing resources (like stone, water, various ores, and tree trunks), building shelters, and avoiding deadly enemies. While game play is open-ended, there are limitations and challenges that influence how the game is played. Creative Mode dispenses with the challenges of Survival and instead provides players with unlimited resources and an open canvas on which to build.

Regardless of mode, players can play by themselves (single player) or with others (multiplayer). When playing multiplayer Minecraft, players interact with other players using text chat on a server. The server provides a single instance of a Minecraft world that many people can access at once. Players access multiplayer Minecraft through player-hosted and business-hosted servers. Large servers like Mineplex and Hypixel support thousands of simultaneous players of all ages. Most organisations hosting servers regulate player behaviour through codes of conduct, which define possible infractions and their consequences. These codes of conduct vary widely, as do the cultures of the servers they are designed to support.

A subset of servers is particularly focused on being kid-friendly and safe. These servers include a number of different additional safeguards designed to limit instances of perceived griefing (e.g., destroying another player's building in-game – see also [23]), the use of bad language, child predation, and aggressive behavior. They also tend to use safeguards like whitelisting, active moderation, and plug-ins to limit toxicity and encourage friendly and age-appropriate play (e.g., blocking offensive words in chat). Our research was done on one such server, as described below.

### The Server

The setting for the research presented here was a moderated free Minecraft server run by a non-profit learning organization, here called CC, for young people aged 8-13. Tech-savvy college students from game design, engineering, and computer science programs serve as the moderators on the server. All are experienced Minecraft players. They build and facilitate in-game activities for kids on the server, and are charged with upholding the server's code of conduct. This code defines the server as a respectful, welcoming, inclusive community that does not tolerate griefing and hurtful/disrespectful language or behavior. Moderators have access to admin level privileges, which allow them to view the chat logs of all players, teleport players into special cool down spaces, and reinstate a player's lost inventory, for example. At least 2 counselors are online any time the server is open. In addition to online counselors,

the server is also whitelisted, which means all players have to be approved to play and the server administrator knows who is on the list. Parents submit and verify their child's Minecraft username in order to be added to the list. Whitelisting helps to ensure that the players on the server are all kids, and that if a serious incident arises on the server that parents or caregivers can be brought into the conversation if needed.

### Moderating practices on other Minecraft servers

Moderation practices on public Minecraft servers tend to meet three primary purposes: reducing toxicity, ensuring fair play, and reducing distractions. Practices are guided by rule books that provide detailed lists of infractions and related punishments. Rules on popular servers like Hypixel and Mineplex<sup>1</sup> encourage players to report infractions either in-game or in the forums. Moderators can use automatic bans and immediate muting in response to rule infractions. Servers expect that players are familiar with their punishment guidelines, and moderators act swiftly when infractions occur. Given that thousands of players play on these servers simultaneously, authoritarian approaches to moderation allow a small number of moderators to "keep the peace." Moderation on the CC's server initially took a similar approach, with one main difference: automatic bans were rare. Counselors were trained to give players the benefit of the doubt, given the players' age and relative inexperience. Across all these servers, players most often see moderation as a necessary, and welcome, form of policing.

## APPROACH AND METHODS

Drawing on the analytical lens from learning sciences, the research focused on examining existing moderation practices and norms within the case study Minecraft server. The first author immersed himself in the case study organisation, here called CC, for a 14 months participatory engagement, separated into two overlapping phases:

### *Phase 1: Understanding moderating practices and challenges*

The aim of the first phase was to understand the existing practices that moderators applied to resolve conflicts at that time. To this end, we applied a combination of participant observation, chatlog data analysis, and interviews. The aim was to unpack the role moderators played on the server, as well as start developing an understanding of how such moderating culture might impact the children's perception of interpersonal conflicts, their role in resolving these, and what 'constructive' strategies to resolve these might be (as modelled by the moderators and other members of the community).

To do so, the first author had: *a*) full access to the CC Slack™ group through which the moderation was organised, with the first author participating in it daily; *b*) an account with moderator powers on the server, leading to approximately 30 hours of in-game observations, captured in written and/or typed notes; *c*) access to all chatlog data among the players and moderator interventions, including 9 months of server history prior to the start of participatory observations; *d*) the history of internal CC team meeting notes; and *e*) weekly meetings with the Head of Program, who was also responsible for leading the moderators training.

<sup>1</sup>E.g., <http://www.mineplex.com/rules>

After 4 months of such participant observation, we conducted 13 formal semi-structured interviews, each approximately 45 minutes long: 9 with Lead Moderators (out of 13 active), 4 with volunteer moderators (out of 6 active), and a final one with the server administrator (who oversaw deployment of all plugins). We also qualitatively analysed 45 days of logs from summer 2015 (30 logs) and April/May 2016 (15 logs).

### Phase 2: Identifying alternative moderation models

Observations in Phase 1 suggested that many of the moderating practices might hinder rather than support children's conflict resolution strategies. To start exploring alternative moderation models, Phase 2 focused on reviewing existing evidence-based conflict resolution programs—originating within Prevention Science—that have been developed for deployment in face-to-face communities (such as high-risk schools), and have been shown to be effective in multiple randomised controlled trials (cf., [2, 22, 34, 44, 53]). Our aim was to articulate a set of underlying principles consistent across the reviewed programs and highlight the associated shifts in moderating culture these would imply for the online moderating practices we observed.

### Analysis

All collected data from Phases 1 and 2 underwent a two-stage thematic analysis process, whereby the data of each phase was at first analysed individually, and then revisited as a whole once the data collection was completed; following the approach outlined by Braun and Clarke [15]. To this end, two of the researchers closely familiarized themselves with the data to identify and systematically search for (recurring) themes. Our findings present the key themes that evolved through this analysis – triangulating the observed daily practices, interviews, informal discussions with the moderators, and iterative engagement with a subset of logs (printed out, highlighted, with margin notes). The articulation of the specific concepts arose from thematic analysis across these diverse data over time (captured in researcher diary and Evernote). It was continuously validated by interaction with the moderating team. We used a similar thematic analysis approach to interpret the review and analysis of the Prevention Science literature and articulate the shifts in moderating practices. The analysis of the log data was used as the last step to further triangulate results across Phases 1 & 2, applying the prevention science framework to a larger set of conflicts. We used Dedoose software to facilitate such log analysis, with the resulting database containing 224 identified conflicts.

### Ethics

The study was submitted through the required ethics committee processes at the host university. All parents of children on CC server gave permission as part of the sign-up process for research observation and server log access. The permission however did not allow for involving children directly, such as through interviews or experiments. All the moderators involved gave explicit informed consent prior to interviews.

## PART 1: UNDERSTANDING THE MODERATION PROCESS

This section presents a discussion of the current moderating process around conflict resolution, building on the analysis of the notes from participant observation, analysis of logs prior to the start of the project (summer 2015 and May 2016), as

well as the 13 interviews with lead and volunteer moderators. In outlining the findings, the rest of this section will emphasise two key themes that served as the key challenges driving research in Phase 2.

The first theme relates to the *roles that moderators self-identified with*: in trying to 'keep peace' on the server, the moderators perceived themselves responsible for establishing the 'truth' of what happened and providing players with resolutions (including punishments). A side effect of the associated moderating practices was a tendency of moderators to monopolise the conflict resolution process, placing the players in the position of passive respondents rather than active agents around any conflicts that arose. Second, we saw the moderating culture as strongly relying on *technological in-game 'solutions'* that strived to reduce—or ideally prevent—the possibility of conflicts altogether. The aim was to prevent children from being able to engage in problematic behaviours at all or quickly roll-back the outcomes if something did happen. This approach kept game play running smoothly by minimising immediate disruption, but failed to support a more empowering set of social norms around conflict resolution among players: a 'solution' to a conflict was to call the moderator. The rest of this section outlines these two themes in detail before we illustrate these observations by selected log examples.

### The role of the moderator: arbiter, teacher, protector

The server we studied was designed to facilitate a 'friendly and safe' online space where kids can pursue an interest in Minecraft in a supportive, social setting. Within this lens, in-game conflicts were implicitly cast as a 'nuisance', as something disruptive to the players and to the culture of the server, potentially leading to negative experiences for kids. As such, moderators were tasked with monitoring the server for conflicts and dealing with them quickly and efficiently, preferably 'out of sight' of the other kids. Such expectations for the moderators' role are inscribed in the Code of Conduct that kids have to accept in order to play on the server. For example, the first rule for players is to "*be nice and respectful to one another*" and that "*ruining another player's experience is not allowed*"; however, if this is violated "*the moderators will step in to resolve the issue*". The next three subsections further unpack the facets of the moderating practices and beliefs that emerged from the interviews, observations, and logs.

### Moderators as arbiters

In line with above, the resonating theme across all of the moderators' interviews was that an important part of their role was to make a decision about what happened and to identify a 'right' response to the situation. In making up their mind about what 'really happened', the moderators often relied on the in-game tools for "*proof*": as M9 describes, in his view "*a lot of times the cases are kind of obvious, as in 'this person ended up grieving this house'*" because that is something he can "*identify through a log-block and I then have a one-on-one conversation with that person*". In such a case, the moderator would review the information from the log block to make a decision on who was the griefer. A subset of moderators described a less mechanistic approach when describing how they go about resolving a conflict: for example, M10 would "*first talk with the person who was affected, then go to the*

*other person to understand what they thought and then kind of pieced the picture together before I engaged someone like ‘you did this’.*

Across all these examples the aim was still to collect enough information on what had happened to make a judgment and assign ‘blame’ (e.g., the log block shows you grieved that house). As we will discuss in more depth, such reliance on in-game logs highlights the emphasis on actions and outcomes (and rolling back to a prior state) rather than either the emotions or motivations behind the behaviours leading to the outcomes.

#### *Moderators as teachers*

A second related thread noticeable across all interviews was the belief that most kids are not malicious and that when they do get into conflict, they often *“just didn’t know better”*. Many disputes between players are due to misunderstandings, differences in perspective, and inexperience. This is especially true of kids new to a multiplayer server, where they can accidentally grief another player simply because they don’t understand the rules and norms of play in a collaborative setting works. It is then the moderator’s role to use their position of authority to get the players to understand what has happened, why it was wrong, and what the ‘right’ behaviour would be. For example, M9 explains how he likes *“to address the issue at hand and explain [...] you’ve been put into cooldown for stealing, you’ve been asked many times to stop, now we’re going to have a chat about it. I try to get them to understand that stealing is not okay and that it affects others. [...] And I finish with like ‘promise me you won’t steal again ... if this comes up again you’ll be facing a hard consequence.”* As we will illustrate with the log examples, this policing approach then manifested in a particular type of authoritarian interaction and ‘voice’ that moderators often used when dealing with conflict.

In line with these observations, moderators felt that their role was crucial for resolving conflicts without escalation that would involve other players. For example, M10 mentions how *“[in any conflict] mostly a moderator needs to be involved. Because sometimes when a kid tries to address it it can be he said she said and players can get upset and that can escalate the situation.”* She continued by reinforcing the authority role of the mod necessary to resolve the situation *“When mods are there, kids respond differently [...] they look up to us so they respond differently than to another kid.”*

#### *Moderators as protectors*

Finally, the moderators felt that their role was to ‘safeguard’ the kids from negative experiences that might be brought about by conflict. For example, an instantiation of this was the work done to distance the ‘victims’ from the conflict resolution process. As M1 described, *“[W]henever I know who did it, I don’t tell them who did it. because I don’t want them to have beef with the other players. whenever they know someone stole something, they always message me back later and say ‘are you talking with the person who did it?’ and usually they won’t let up until I actually do it and actually talk with them. If they know I’ve talked with them then they’re ok with it. [...] The older the kids get, the more they are concerned about whether or not there will be justice.”* Even if the ‘victims’ knew the name of whoever stole or grieved their stuff, the approach was often that the mods reinstated what was lost as

quickly as possible: returned the house to the prior un-grieved state using server history, or asked the player what had been stolen from them and gave them copies of the items. Only then would the moderator start discussing the situation with the ‘guilty’ party.

This distancing role was instantiated already in the moderators’ training materials. Within the FAQ for what to do in case of grieving, the process was to: (1) Assess the situation and ask what happened/what got destroyed; (2) If the player knows who ‘did it’, let the player know you will deal with the situation/suspect; (3) Fix the player’s grieved area. Moreover, the players were instructed to ‘notify a mod’ if anything bad has happened, rather than attempting to resolve it themselves.

Overall, this led to a particular culture that nudged the players to ‘call for help’ at the first sign of trouble, and encouraged the moderators to hide the resolution processes behind the scenes; with the ‘physical’ aftermath of the conflict (such as a broken home or lost items) being miraculously resolved through mod in-game powers. We now describe the in-game tools that moderators used to enact these moderating practices.

#### **In-game tools as technological ‘solutions’ to conflict**

In running a Minecraft server, the moderators have access to technological tools that provide, potentially, complete control over their and others’ avatar capabilities in-game: they can teleport people instantly or freeze them in place, they can add items into inventories or clear them away, and they can mute or otherwise constrain what players can say and to whom.

With respect to conflict resolution, the technological tools we saw being used can be divided into three somewhat overlapping categories: (i) those that affect the communication with a child in the aftermath of a conflict (in particular the ‘cooldown space’ described below); (ii) those that help revert changes due to other players behaviours; and (iii) those that affect in-game structures, access to objects, or types of available interaction more broadly, with the aim of reducing conflict overall. We describe each in turn below.

##### *Communication tools: the cooldown space*

Cooldown is the most frequently utilised tool during conflicts. Placing someone in cooldown transports that player into a specific area of the server, where players are stripped of their abilities to build or destroy a block (i.e., they cannot interact with the world or other players in the cooldown space); and also all channels but the one used to communicate with the moderator are muted (i.e., kids are virtually cut away from everything and everyone, including the conflict participants).

Within the server culture, cooldown played a dual function of both a ‘calming down’ as well as a ‘time out’ space. A common narrative among the moderators was that children were being brought into cooldown to ‘explain what they did wrong’ and were not to be let out unless it was clear ‘they understood’. The moderators valued cooldown as a place to discuss things with players ‘without interruptions’. At the same time, they also appreciated that, for many kids, being put into cooldown could feel like being put into prison.

The interactions around the cooldown space illustrate how, although the moderators have complete control *over avatars*

capabilities in-game, they can also lose all control *over the player* if he or she decides to disengage from the conversation or even disconnect from the game. For example, M2 describes the striking differences she sees to her real-world experiences in conflict resolution: “[*What was surprising for me was*] the people’s ability to completely ignore you. they just don’t type back to you. This can’t happen in the real space, where they stand right in front of you. In the real world, they [can] pretend they don’t hear you, cross their arms ... but in digital space you have no feedback at all.” Similarly, other mods emphasised how important it was to make sure kids “don’t feel like they are trapped in a situation .. because a lot of times the kids would just log off the server at that point. ”

#### *Roll-back tools: re-instating status quo*

While cooldown was used to make the ‘perpetrator understand’ what they did ‘wrong’, a second set of tools was used to ‘resolve’ conflicts by rolling back the state of the game to the moment prior to the conflict. The moderators would use the *log block* tool to determine who touched blocks in a particular area and when (to determine ‘griefing’) as well as to see who took items from a chest (to determine ‘stealing’). Similarly, they would use the *invsee* command to look into and manipulate players’ inventories, for example to commandeer items deemed as being stolen. Such reliance on *log-block* or *invsee* as an objective indicator of ‘truth’ put emphasis on whether an action had happened (did he break the house), rather than what the player was thinking at that point, i.e., the intentions behind the actions (was he being malicious, thinking he was helping, or was it a mistake?).

Moderators used the same set of tools to resolve the conflict for the ‘victims’: *log block* allowed them to restore buildings to a previous point in history (rolling back the changes), and they placed new copies of items into the victim’s inventory through *invsee*. We often saw moderators tell players to “not to worry, tell me what you lost and I’ll replace your items”, as a way of trying to quickly erase the aftermath of escalating conflicts. M8’s quote below illustrates these goals of dissolving any emotional ‘tension’: if “*somebody got something stolen and they were getting really... like, riled up about it. Like really frustrated so I’m like “You know what? Just tell me what the enchantment was or whatever and I’ll remake the sword or whatever” so like... you know, so that the conflict doesn’t build up and get them angry*” A related strategy was referred to as the ‘distraction method’: the aim was to distract the child with another activity to help them ‘forget about’ the issue and let emotions ‘sizzle out’.

The moderators seemed to be implicitly distinguishing between what we will call *mechanical* and *player* conflicts, drawing on the (limits of the) tools they had at their disposal: As M9 mentioned, “*a lot of times the cases are kind of obvious [and] we can identify that through a log-block and I then have a one-on-one conversation with that person. Obviously it is very different to having two people in conflict with each other where it is not such a cut-and-dry case of who was wrong.*” In other words, the *mechanical* conflicts are those where an ‘objective’ interpretation can be easily determined and rolled back by technological tools, while *player* conflicts—such as killing another player—were seen as messy emotional situations that are not possible to roll back so easily. So although

killing and stealing might lead to similar negative emotions for the person involved, the moderating approach saw these differently: Stealing was considered to be ‘solved’ by the mod returning copies of the items and it was thus perceived as ‘easy’ to be resolved. In contrast, the moderators considered killing—which usually did not lead to any actual loss of items and the avatar got ‘reborn’ in seconds—much harder to deal with as there was nothing to mechanically ‘roll-back’.

The moderators’ perception of ‘easy resolution’ within mechanical conflicts was however not consistent with the log data or moderators’ own descriptions: Recall, for example, the length to which moderators went to not tell the player who stole their items (so as to avoid expected emotional outbursts); or the example from the logs in the following section, where players were clearly not content even after being reimbursed. Overall, while the mechanistic roll-back was important in restoring the status quo within the game and a key part of moderation, it often still did not address the emotional facet of the conflict and the perceived need for ‘justice’.

#### *Structural changes to reduce conflicts*

From the start of running the server, and based on seeing conflicts arise, CC continuously worked on iteratively re-designing the structure of the Minecraft world and activities so as to reduce the likelihood of conflicts appearing altogether. We note that the moderating culture described above developed in parallel to these structural changes, as a way to deal with conflicts that still appeared.

The uniting feature of all the developments described here is the focus on limiting capabilities of players’ avatars so that known problematic behaviours would no longer be possible. For example, to deal with an increase in stealing from chests, children were instructed (and given) ‘ender chests’ as a specific type of in-game container that could be opened only by the owner. Similarly, attempts to reduce griefing relied on using ‘claims’, i.e., tools with which both kids and moderators could claim a part of the world so that again only they (and other players they explicitly invite) could modify blocks in that area.

However, conflicts still arose despite all these additional safeguards that were iteratively introduced and the energy required from moderators to deploy them. As M12 put it, there are two issues: “[*First*] the grief prevention only works, if used correctly, but it relies really on the kids and counselors. [*Second*] it doesn’t solve the underlying problem of someone wanting to do something they shouldn’t. It doesn’t resolve the hard issues, it just keep them from being able to commit those actions in that particular space.” As an example, we saw instances of ‘griefing with anti-grief tools’, such as someone claiming a part of another person’s build if they noticed the claims weren’t there. Similarly, even if moderators made sure everything was covered by claims and ender chests, they still could not prevent the killing of other people’s pets and this resulted in a stable stream of conflict instances. And finally, limited capacity of ender chests made them un-scalable for deployment in larger group situations, bringing back the need for normal chests and issues with stealing associated with these.

Overall, these attempts seemed to further emphasise the limits of the focus on technological solutions (through prevent-

ing/rolling back problematic behaviours) and reliance on moderators' solutions, rather than focusing on developing more empowering norms to guide kid's actions in the first place.

### Illustrating these themes within logs

In this section, we select snippets from two conflicts that happened on the server. Taken together, they illustrate many of the themes we discussed above: the moderators arbitrate, try to teach, as well as 'protect' the kids from each other by resolving the conflict separately; they work within cooldown space to make this separation possible, as well as try to resolve the conflict by rolling-back the 'damage caused'.

#### Example 1

In this example, a player called P1 died because another player P2 pushed her (accidentally) into lava. This is a problem as P1 lost all her equipment and items she might have been collecting for some time (e.g., an hour or longer).

[17:17:53] P1: what the heck!!!!!!!!!!!!!!!!!!!!!!  
 [17:17:57] P1: why did you do that!!!!!!!!!!  
 [17:18:03] P1: i am so freaking mad at you!!!!!!  
 [17:18:16] P1: why the bloody heck would you do that?!?!?!?!?!?!?!?!?!?!  
 [17:18:20] P2: i was nerbous that uou were acting weird?  
 [17:18:24] P1: why the heck!!!!!!!!!!  
 [17:18:33] P1: how the Heck am i acting weird?!?!?!?!?!  
 [17:18:36] P2: sorry  
 [17:18:36] P1: I need a mod!!  
 [17:18:39] P1: right now!!!

By this time, the moderators appear, offering to compensate P1 for any items she lost:

[17:19:38] Mod1: we can recomp you  
 [17:19:50] P2: i wanted her to be safe, and i got nervois when she froze  
 [17:19:51] P1: and he freaking dropped me in lava?!?!?!?!?!?!?!?!  
 [17:19:53] Mod1: P1 go home and i'll get your stuff back  
 [17:19:57] P1: why the bloody heck!!!!!!!!!!!!!!!!!!!!!!  
 [17:20:06] P1: wahhhhhH!!!!!!!!!! ¿:((((  
 [17:20:18] P1: and i had freaking 20 levels!!!!!!!!!!!!!!!!!!!!!!!!!!!!  
 [17:20:24] P1: why did he do that?!?!?!?!?!?!?!?!?!?!  
 [17:20:28] Mod1: P1, go home and i'll give you your stuff back

When it is clear that this offer is not working—after twelve more messages from P1 showing her frustration within two minutes—the moderator teleports P1 to himself and enables a TNT explosive, which is otherwise disabled on the server. They then go on a destruction spree.

[17:21:28] Mod1: lets go to stress relief  
 [17:24:04] P1: you're right this is fun  
 [17:24:08] P1: i like blowing stuff up  
 [17:24:13] P1: it gets out my anger  
 [17:24:15] P1: ¿:(  
 [17:25:08] Mod1: lol only if it makes you kind of feel better  
 [17:25:12] Mod1: but only within reason

During all this time, P2 is trying to apologise and make up with P1; however, he is ignored by both P1 and the mods.

[17:21:45] P2: P1 you can loot my house  
 [17:22:18] P2: how about we do something else to calm you down?  
 [17:23:12] P2: sorry  
 [17:23:31] P2: sorry  
 [17:23:45] P2: i wont do it again  
 [17:24:47] P2: want to explore with me?  
 [17:24:53] P2: i promise  
 [17:25:06] P2: im sorry P1  
 [17:25:43] P2: sorry P1

[17:25:59] P2: accept my apology?  
 [17:26:03] P1: no  
 [17:26:28] Mod1 → P2: its okay, we're trying to calm her down

#### Example 2

In this example, two friends are arguing because one destroyed the house of another. The moderators quickly intervene, the kids are moved to the cooldown space, and talked to separately by different moderators. Notice how both moderators are attempting to 'teach' P4 that what they did was wrong, but this elicits only very limited responses from the player.

[19:31:17] Mod2: P4  
 [19:31:22] Mod2: why did you destroy P3's stuff?  
 [19:31:37] P4: i did not  
 [19:31:42] Mod2: yes P4  
 [19:31:44] Mod2: yes you did  
 [19:31:47] Mod2: I checked the logs  
 [19:31:49] Mod2: they said you did  
 [19:31:56] Mod2: don't try to lie please

— no response from the child for 8 minutes —

[19:40:19] Mod3: Hi P4, do you want to come talk this out?  
 [19:41:50] P4: yes  
 [19:42:25] P4: yes  
 [19:42:37] Mod3: So you destroyed parts of P3's house, and I know you're near a ban  
 [19:43:51] Mod3: Remember, destroying other kids' houses isn't allowed, and you need to follow the code of conduct for camp, or you'll have to deal with the consequences :(  
 [19:44:42] P4: ok

## PART 2: ALTERNATIVE MODERATING MODELS

While the technological tools and existing moderating practices can be seen as efficient at keeping peace in the short term, we can also characterise these as engendering an authoritarian/policing stance (cf., [9, 21]). A consequence of these tools and practices, we argue, is that they also bring substantial implications for how the children's role is positioned and their subsequent learning opportunities.

In particular, by emphasising the role of the moderator as an arbiter to establish the 'truth', roll back the 'aftermath' of behaviour, and mete out 'learning' (through punishment if need be), the moderating practices seemed to lead to:

- the 'victims' being excluded from the resolution process ('restoring the world' should make things okay)
- the 'perpetrators' being often unable to 'fix things' themselves (as the changes have been rolled back by the mods ... what action could they use to show they are sorry?); and
- the moderators' intervention mostly not addressing underlying interpersonal emotional/motivation/behavioural issues.

We argue that such a moderating culture misses out on opportunities to empower children to develop their own conflict resolution skills as part of their in-game engagements.

In what follows, we draw on existing literature of in-school conflict resolution interventions originating in Prevention Science to offer a conceptual grounding for an alternative, competency building approach to conflict resolution. In doing so, we first briefly review the Prevention Science literature. We then draw out a particular typology of conflict mediation which can orient designers to specific changes on the level of moderators' utterances (and training) to engender more empowering mediating culture. Finally, we build on this framework to articulate a set of underlying principles consistent across the

	Power	Suggestion	Explanation	Problem solving
<b>Assumed mechanic of change</b>	Influencing the children's behaviour by 'being strong and imposing their will', such as by yelling, belittling, and (physical) punishment	Telling children how to solve the situation by providing possible solutions.	Providing solutions, followed with an explanation of why this matters/works.	Involving children in thinking about what they are doing and why, with the aim that they reach acceptable solution.
<b>Example from ICPS</b>	"How many times I have to tell you ... !"; "Stop that yelling right now, or you'll face consequences!"	"Why don't you do ... ?"; "Just ask him to leave you alone, but do it quietly."	"Don't do this, it makes others uncomfortable."; "If you don't stop yelling, you won't have any friends."	"What do you think the other kids might say or do after you yell at them?" "Can you think of a different way to let them know how you feel?"
<b>Example from Minecraft</b>	Do you hear me? Stop breaking those blocks or you'll be in cooldown.	You shouldn't be destroying those blocks. Can you stop it now please?"	You shouldn't be destroying those blocks. The person who built that will feel sad."	How would you feel if someone destroyed your house? What else could you do instead?

Figure 1. Four types of conflict resolution as outlined in ICPS.

multiple reviewed programs and highlight the associated shifts in moderating culture these would imply for the online moderating practices we observed. Our hope is that this conceptual model can inspire and guide other designers and researchers to develop a more empowering moderating systems.

### Prevention science programs in education

Prevention Science is a discipline that examines how to prevent or moderate behavioural and/or mental health issues through large scale interventions, evaluated in real-world deployments [18, 19, 22, 28, 29, 65]. For example, so called 'social-emotional learning' (SEL) curricula are now in more than 44% of US schools [16], with positive immediate and long-term effects shown in randomised controlled trials [3, 19, 22, 64]. Slovak and Fitzpatrick [54] and Conley et al. [19] provide recent reviews across both HCI and Prevention Science literatures.

A subset of such Prevention Science programs—such as RCCP [20], Peacemakers [33], I Can Problem Solve [51]—are specifically designed to develop students' conflict resolution competencies; and have been implemented in varied face-to-face communities such as high-risk schools or whole US school districts with evidence of reduced peer-conflicts [2, 34, 44] and bullying [53]. It is these specialised conflict resolution programs that we are referring to in what follows, starting with a specific framework that can illustrate how such alternative conflict resolution approaches might look.

### Conceptual framework: Four types of conflict mediation

Originating within the "I Can Problem Solve" (ICPS) program [18, 51–53], the ICPS framework is well positioned to inform online moderation practices as it considers ways in which those in a position of 'power'—whether that is power status associated with being a parent, or that of a moderator on an online server—can intervene and mediate when conflicts arise. It distinguishes four ways of approaching such mediation: power, suggestion, explanation, and problem solving. See Figure 1 for descriptions of each of the types.

As with all typologies, the four types of conflict resolution framework present a highly simplified version of reality. However, such simplification can often serve as a useful sensitising concept [13], providing designers and moderators with a set of metaphors and tools-to-think with. We argue that what makes the ICPS framework particularly appealing in this regard is that it can be directly applied to individual statements that the moderator makes during the conflict resolution; while still being based on a holistic and evidence-based approach. As such,

it not only suggests a sample language of how moderators 'should' be resolving conflicts, but also provides a conceptual model to identify and troubleshoot problems in any given moderating culture.

To illustrate this, we returned to our log data and used the ICPS typology to code a random sample of conflicts identified in 30 days of server logs<sup>2</sup>. The results showed that power was by far the most used conflict resolution approach (69% of conflict instances), followed by suggestion (14%), explanation (13%) and only a handful of problem solving resolution (4%). Beyond corroborating the qualitative results presented in Part 1, this conceptual framing then helped us in identifying the types of statements responsible for the observed distribution of conflict mediation styles, how these should be altered to move the moderating practices toward problem solving, and also how to communicate this intended shift in moderating culture to the moderators.

### Principles and associated shifts in perspective

Given the existing Prevention Science evidence base, and comparing it to the findings previously presented about current moderation practices, we articulate the principles through which the face-to-face SEL programs attempt to provide students with opportunities to learn the necessary skills themselves. We then frame this as a suggested set of 'shifts in perspective' when contrasted to the more traditional discipline programs (and approaches such as those we saw in online Minecraft moderation). See Figure 2 for an overview.

#### Principle 1: From a 'nuisance' to a 'learning opportunity'

The underlying philosophy across all of these programs is that conflicts are a normal and inevitable part of daily life that cannot be eliminated. It is rather how conflicts are managed—not their presence as such—that determines if they are destructive or constructive [2, 20, 34, 35, 53]. In fact, since experiential learning is necessary to develop constructive ways of responding to conflicts [6, 54] it is actually crucial for children to have experiences of conflict. However, an important role of these programs is to ensure that such experiences happen only in a 'safe space', where escalation does not lead to serious outcomes (e.g., physical violence), and where support is available for students to try out constructive rather than destructive conflict resolution methods.

<sup>2</sup>We sampled 30 days of logs randomly selected from Jul-Aug 2015. This led to a sample of 197 conflict instances that included sufficient moderator involvement to be coded. First and third author coded the data, with any differences in coding resolved by discussion.

Principle	Associated shift in perspective	Level of application
Conflict are normal and inevitable: it is how they are managed what makes them constructive or destructive.	From nuisance to learning opportunity	Server culture
People learn by resolving conflicts themselves rather than by having been provided solutions by others.	From 'telling' to 'asking'	Moderator role
Truth of what happened is less important than what can we do about it now.	From 'assigning blame' to 'finding solutions'	Aim of specific mediation
Emotions are at the core of most conflicts; reverting state does not resolve feelings.	From 'repairing things' to 'resolving feelings'	Aim of specific mediation

**Figure 2. Principles and associated shifts in perspective underlying conflict resolution programs in Prevention Science.**

The first principle is then re-framing conflicts from something to be avoided at all costs to a potential learning opportunity: SEL programs see conflicts as valuable teachable moments and it is the role of the (moderation) programs to provide and facilitate a safe space for learning from conflicts: engendering a positive set of norms and mindset around conflict, offering resources for students, and scaffolding the development of budding skills through in-person support.

#### *Principle 2: From 'telling' to 'asking'*

The second shift in perspective refers to the role of the mediator: all the programs we reviewed emphasise how the traditional, authoritarian approaches teach students that external figures are needed to resolve conflict, thus hindering their own competency development. In contrast, the mediator in SEL conflict resolution programs is expected to refrain from giving solutions and instead facilitate decision making through a series of open ended prompts. The emphasis is on scaffolding the children's thinking processes through a carefully designed dialogue, rather than telling them the solution. This involves a feeling of a 'loss of control' over the conversation for the mediator; however, this is also an indicator of the empowerment for the child, which is necessary for the learning to take place.

The curricula have developed and validated language and tools to help mediators guide the process in ways that are more likely to facilitate constructive resolution process. For example, if a solution seems clearly inappropriate ('I will hit him back'), rather than rejecting such a solution from a position of power, the program might suggest that the adult either probes for an alternative solution ('that is one option, what else could you do?'), guides the child to think through the consequences ('how do you think he will react then?'), or asks the other child whether that would resolve the situation successfully from his/her perspective ('would this solve the problem?').

#### *Principle 3: From 'assigning blame' to 'finding solutions'*

Third, understanding 'the truth' about what happened with the associated blame and retribution are no longer seen as the ultimate goal of mediation; and, in fact, not even seen as possible. The SEL curricula emphasise that perceptions of

'what happened' might be very different across those involved in a conflict. Instead of dwelling on the past, the curricula suggest outlining each person's perspective and then quickly moving towards clarifying what would make the situation better for each of the participants now ('what do you want to happen?'), and what could be mutually acceptable solutions ('what can we do to resolve this?').

Aligned with the 'asking not telling' approach, it is particularly important that those involved in the conflict are able to acknowledge each others' perspectives on what has happened (even if these accounts are different from each other), rather than the mediator finding the single 'truth'. It is the mediators' role to facilitate this process, helping those in conflict spend most energy and time on developing possible solutions rather than fighting over 'what happened'.

#### *Principle 4: From 'repairing things' to 'resolving feelings'*

Finally, all the curricula emphasise the importance of addressing feelings during the conflict resolution and mediation process; these feelings are seen at the heart of most conflicts. In essence, conflicts are not seen as a direct result of an action but rather of the *interpretation* of that action: For example, if you thought someone bumped into you on purpose to spill your coffee, you would likely feel differently than if you thought it was an accident. You would also likely have different ideas about what might be a solution. In the latter example, a refilled cup of coffee might be sufficient; in the former, coffee refill would unlikely resolve the anger you might feel towards the other person. This exemplifies the importance of perceived intent and an innate need for 'justice', underpinning many conflicts. It also highlights that just reverting any physical outcomes of conflict to a previous state – such as refilling coffee or reinstating a minecraft build – does not necessarily resolve the emotional turmoil that drives the conflict.

#### *Summary*

As outlined in Figure 1, the shifts in perspective can be seen as operating on different levels of interaction: Starting at the overall culture and norms around conflict that are developed and reproduced within social groups (Principle 1); continuing to the roles that moderators see themselves as playing within conflict resolution generally (Principle 2); as well as the approach they should take within individual conflict instances (Principles 3+4). The next section moves to start unpacking the implications that acceptance of these principles would have for online moderation—both in Minecraft as well as other online spaces for youth—as well as scopes when and how these approaches might be applicable.

## DISCUSSION

The principles outlined above provide a striking contrast to the practices we observed within the Minecraft moderating culture. The competency building literature reviewed in Part 2 has further highlighted how such moderation practices placed the players in the position of passive respondents rather than active agents around any conflicts that arose. Moreover, these practices did not lead to a sustainable or stable solutions: individual issues may have been resolved quickly but players often repeated the behaviors soon after; and the 'solution' to a conflict was to call the moderator. So while the punitive approach might have brought short-term benefits of 'quick'

resolutions minimising immediate disruption, it also left the burden of the conflict resolution on moderators only.

### Implications for online moderation

We saw the hindering or supportive effects of moderation arise from the combination of (i) the specific support tools available for moderators and the players in-game; (ii) the role that moderators strive to enact; and (iii) the associated server culture and established norms. We argue that the design of the existing ‘anti-griefing’ tools seems to have played a strong role in enabling and reproducing the authoritarian approach to—and culture of—moderation (cf., [9, 21]). It was through the tools such as log-block that the moderators saw themselves as able to ‘uncover the truth of what happened’; and tools like cooldown that allowed for the complex set of practices through which moderators monopolised the conflict resolution process. Similarly, the ability to manipulate the text-based communication channels at will gave moderators the power to effectively distance the children from decision making; albeit under the well-intended goal to limit any negative experience that might arise from the conflict for the ‘victim’.

While it is still an open question how the conflict resolution strategies and language from Prevention Science SEL programs can be re-appropriated into an online space, the principles and associated shifts in perspective outlined above can provide an initial set of design implications to guide exploration of such player-empowering socio-technical moderating structures. More broadly, by highlighting the possible positive role of conflict for competency development, the SEL literature does not argue for removal of moderated environments, but rather suggests the possibility of empowering (rather than authoritarian) moderation approaches.

In fact, we propose that the very affordances of the game-space that made the authoritarian approaches possible might instead be utilised in the service of empowering children to learn. For example, the ability to fully control the communication channels can be used to not only punish or separate players, but also to provide contextualised scaffolding and support for children to voice disagreements in non-escalatory ways. The existing social-emotional programs provide examples of well thought-through language and ‘scripts’ to promote constructive disagreements that, however, rely on extensive training and memorisation before children are able to use these in actual conflict. An in-game contextualised scaffolding could simply side-step this issue, offering the players access to the conflict resolution language and process as they type, as well as automatically flagging framings that are known to be escalatory (e.g., “you never/always ...”). Moreover, the in-game interactions provide a steady stream of meaningful emotional engagements (often including peer-disagreements or conflict) that could be appropriated as teachable moments in ways that are impossible in ‘physical-world’ settings: the conflicts can be made available for post-hoc reflection and coaching as all interactions immediately transcribed in logs, and could also be deliberately designed for through in-game activities (cf., [55]).

Together, these observations point to the possibility of utilising such online spaces for a novel kind of ‘situated SEL curricula’, where students are scaffolded in learning social-emotional competencies from meaningful interpersonal situa-

tions they naturally encounter. Such a vision of making online spaces ‘safe through learning’ strives to create server culture—and tools—that empowers players’ agency and competencies when conflict arise, while keeping the moderating oversight to step in should events escalate. As such, it would not remove the responsibility for keeping children safe, but rather suggest alternative safeguarding mechanisms to the externally imposed restriction and regulation exemplified elsewhere (e.g., the adult-based oversight outlined in Autcraft [46], muting and vaporising in Whyville [37], or the pre-typed response trees in WebKinz [11]).

### Broader applicability

The intricate moderating structures described above are not a silver bullet against encountering toxicity online; and are unlikely to fit into every online community. However, they can orient us to the potential of developing social competencies through online gameplay, uncover the moderating mechanisms that support/limit this learning, and start envisioning how supportive moderating cultures could be designed for.

To quote Amy Bruckman again, “*social solutions require time, effort, and leadership [...] and having time to engage individual players in dialogue is a luxury*” [17]. On a pragmatic level, we expect it will be harder to implement empowering learning structures into the semi-anonymous, hyper competitive setting of League of Legends, rather than into often tightly-knit online tween communities. Moreover, across any online setting, including such learning agenda will likely surface tensions around the varied perspectives of what types of experiences online gaming spaces should support—from a fun place to relax (players), to safety first (parents), to learning (educators), to a smoothly working community (moderators)—and thus also where moderators’ energy should be used, and what in-game technology support should be provided.

To draw on experiences from in-school programs, it is likely that these tensions would stem from a clash between long- and short-term effects: The in-school programs require an initial intensive training phase to change social norms and train peer-mediators, but then are self-sustainable through emergent peer-mediation practices. Similarly, in shifting to an empowering moderating culture, we expect each conflict instance might initially require comparatively more moderator effort to resolve; but the overall moderating effort would progressively diminish as players learn to resolve conflicts and social norms around constructive resolution process develop.

### CONCLUSIONS

This paper presents findings from a long term participatory engagement with a moderated Minecraft server for young gamers (aged 8-13). By combining such ethnographic case study data with scholarship from learning sciences, we argue that the existing authoritarian moderating practices can not only prevent young gamers from actively engaging with and learning from conflicts they experience online; but can also reinforce a disempowering set of social behaviours and values more broadly. As a first step towards the design of alternative moderating models, we draw on Prevention Science interventions to articulate a set of principles to guide future designs.

## REFERENCES

1. 2015. *Essential Facts about the Computer and Video Games Industry (Annual report)*. Technical Report. Association, Entertainment Software. <http://www.theesa.com/wp-content/uploads/2015/04/ESA-Essential-Facts-2015.pdf>
2. J. Lawrence Aber, Joshua L. Brown, and Stephanie M. Jones. 2003. Developmental trajectories toward violence in middle childhood: Course, demographic differences, and response to school-based intervention. *Developmental Psychology* 39, 2 (2003), 324–348. DOI : <http://dx.doi.org/10.1037/0012-1649.39.2.324>
3. Yaser Adi, Amanda Killoran, Schrader McMillan, Amanda Kiloran, and S Steward-Brown. 2007. *Systematic review of the effectiveness of interventions to promote mental wellbeing in children in primary education. Report 1: Universal approaches: non-violence related outcomes*. Technical Report June 2007. National Institute of Health and Clinical Excellence Report (NICE).
4. Diego Fernando Gutierrez Aponte and Deborah Richards. 2013. Managing cyber-bullying in online educational virtual worlds. In *Interactive Entertainment '13*. ACM Press, New York, New York, USA, 1–9. DOI : <http://dx.doi.org/10.1145/2513002.2513006>
5. M. E. Ballard and K. M. Welch. 2015. Virtual Warfare: Cyberbullying and Cyber-Victimization in MMOG Play. *Games and Culture* (jun 2015). DOI : <http://dx.doi.org/10.1177/1555412015592473>
6. R Bar-On, K Maree, and MJ Elias. 2007. *Educating people to be emotionally intelligent*. Greenwood Publishing Group.
7. Sasha a Barab, M Gresalfi, and Anna Arici. 2009. Why educators should care about games. *Educational Leadership* 67, 1 (2009), 76–80. DOI : <http://dx.doi.org/10.1017/CBO9781107415324.004>
8. S. A. Barab, M. Gresalfi, and A. Ingram-Goble. 2010. Transformational Play: Using Games to Position Person, Content, and Context. *Educational Researcher* 39, 7 (2010), 525–536. DOI : <http://dx.doi.org/10.3102/0013189X10386593>
9. Diana Baumrind. 1991. The influence of parenting style on adolescent competence and substance use. *The Journal of Early Adolescence* 11, 1 (1991), 56–95.
10. Max V Birk, Benjamin Buttler, Jason T Bowey, Susanne Poeller, Shelby C Thomson, Nicola Baumann, and Regan L Mandryk. 2016. The Effects of Social Exclusion on Play Experience and Hostile Cognitions in Digital Games. *Chi '16* (2016), 1–14. DOI : <http://dx.doi.org/10.1145/2858036.2858061>
11. Rebecca W Black and Stephanie M Reich. 2012. 14 Culture and Community in a Virtual World for Young Children. *Games, learning, and society: Learning and meaning in the digital age* (2012), 210.
12. Jeremy Blackburn and Haewoon Kwak. 2014. STFU NOOB!. In *Proceedings of the 23rd international conference on World wide web - WWW '14*. ACM Press, New York, New York, USA, 877–888. DOI : <http://dx.doi.org/10.1145/2566486.2567987>
13. H. Blumer. 1954. What is Wrong with Social Theory? *American Sociological Review* 19, 1 (1954), 3–10. DOI : <http://dx.doi.org/10.2307/2088165>
14. Danah Boyd. 2014. *It's Complicated: the social lives of networked teens*. Yale University Press.
15. Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101. <http://www.tandfonline.com/doi/abs/10.1191/1478088706qp063oa>
16. J Bridgeland, M Bruce, and A Hariharan. 2013. The missing piece: A national teacher survey on how social and emotional learning can empower children and transform schools. (2013).
17. Amy Bruckman, Pavel Curtis, Cliff Figallo, and Brenda Laurel. 1994. Approaches to managing deviant behavior in virtual communities. In *CHI Conference Companion*. 183–184.
18. John D Coie, Norman F Watt, Stephen G West, J David Hawkins, Joan R Asarnow, Howard J Markman, Sharon L Ramey, Myrna B Shure, and Beverly Long. 1993. The science of prevention: A conceptual framework and some directions for a national research program. *American psychologist* 48, 10 (1993), 1013.
19. Colleen S. Conley, Joseph A. Durlak, Jenna B. Shapiro, Alexandra C. Kirsch, and Evan Zahniser. 2016. A Meta-Analysis of the Impact of Universal and Indicated Preventive Technology-Delivered Interventions for Higher Education Students. *Prevention Science* 17, 6 (2016), 659–678. DOI : <http://dx.doi.org/10.1007/s11121-016-0662-3>
20. William DeJong. 1994. *Building the Peace: The resolving Conflict Creatively Program*. <http://www.ncjrs.gov/App/abstractdb/AbstractDBDetails.aspx?id=149549>
21. Sanford M Dornbusch, Philip L Ritter, P Herbert Leiderman, Donald F Roberts, and Michael J Fraleigh. 1987. The relation of parenting style to adolescent school performance. *Child development* (1987), 1244–1257.
22. Joseph A Durlak, Roger P Weissberg, Allison B Dymnicki, Rebecca D Taylor, and Kriston B Schellinger. 2011. The impact of enhancing students' social and emotional learning: a meta-analysis of school-based universal interventions. *Child development* 82, 1 (2011), 405–32. DOI : <http://dx.doi.org/10.1111/j.1467-8624.2010.01564.x>
23. Cy Foo and Emi Koivisto. 2004. Defining grief play in MMORPGs: Player and developer perceptions. *Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology* (2004), 245–250. DOI : <http://dx.doi.org/10.1145/1067343.1067375>

24. Colin Gallagher. 2014. *An Educator's Guide to Using Minecraft® in the Classroom: Ideas, inspiration, and student projects for teachers*. Peachpit Press.
25. James Paul Gee. 2008. *What video games have to teach us about learning and literacy*. Macmillan. 256 pages.
26. Douglas A. Gentile, Craig A. Anderson, Shintaro Yukawa, Nobuko Ihori, Muniba Saleem, Lim Kam Ming, Akiko Shibuya, Albert K. Liau, Angeline Khoo, Brad J. Bushman, L. Rowell Huesmann, and Akira Sakamoto. 2009. The Effects of Prosocial Video Games on Prosocial Behaviors: International Evidence From Correlational, Longitudinal, and Experimental Studies. *Personality and Social Psychology Bulletin* 35, 6 (2009), 752–763. DOI : <http://dx.doi.org/10.1177/0146167209333045>
27. Isabela Granic, Adam Lobel, and Rutger C M E Engels. 2014. The benefits of playing video games. *American Psychologist* 69, 1 (2014), 66.
28. Mark T Greenberg. 2006. Promoting resilience in children and youth: preventive interventions and their interface with neuroscience. *Annals of the New York Academy of Sciences* 1094 (dec 2006), 139–50. DOI : <http://dx.doi.org/10.1196/annals.1376.013>
29. Mark T. Greenberg. 2010. Schoolbased prevention: current status and future challenges. *Effective Education* 2, 1 (mar 2010), 27–52. DOI : <http://dx.doi.org/10.1080/19415531003616862>
30. Mark D Griffiths, Mark N O Davies, and Darren Chappell. 2004. Demographic factors and playing variables in online computer gaming. *CyberPsychology & behavior* 7, 4 (2004), 479–487.
31. Mizuko Ito, Sonja Baumer, Matteo Bittanti, Rachel Cody, Becky Herr Stephenson, Heather A Horst, Patricia G Lange, Dilan Mahendran, Katynka Z Mart\`inez, C J Pascoe, and Others. 2009. *Hanging out, messing around, and geeking out: Kids living and learning with new media*. MIT press.
32. Mizuko Ito, Kris Gutiérrez, Sonia Livingstone, Bill Penuel, Jean Rhodes, Katie Salen, Juliet Schor, Julian Sefton-Green, and S Craig Watkins. 2013. *Connected learning: An agenda for research and design*. BookBaby.
33. David W Johnson and Roger T Johnson. 1995. *Reducing school violence through conflict resolution*. ASCD.
34. David W. Johnson and Roger T. Johnson. 1996. Conflict Resolution and Peer Mediation Programs in Elementary and Secondary Schools: A Review of the Research. *Review of Educational Research* 66, 4 (1996), 459–506. DOI : <http://dx.doi.org/10.3102/00346543066004459>
35. Tricia S. Jones. 2004. Conflict resolution education: The field, the findings, and the future. *Conflict Resolution Quarterly* 22, 1 (2004), 233–267. DOI : <http://dx.doi.org/10.1002/crq.100>
36. Yasmin B. Kafai. 2010. World of Whyville. *Games and Culture* 5, 1 (2010), 3–22. DOI : <http://dx.doi.org/10.1177/1555412009351264>
37. YB Yasmin B Kafai and Kristin A Searle. 2010. Safeguarding Play in Virtual Worlds: Designs and Perspectives on Tween Player Participation in Community Management. *International Journal of Learning* 2, 4 (2010), 31–42. DOI : <http://dx.doi.org/10.1162/ijlm>
38. Haewoon Kwak, Jeremy Blackburn, and Seungyeop Han. 2015. Exploring Cyberbullying and Other Toxic Behavior in Team Competition Online Games. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*. ACM Press, New York, New York, USA, 3739–3748. DOI : <http://dx.doi.org/10.1145/2702123.2702529>
39. Jean Lave and Etienne Wenger. 1991. *Situated learning: Legitimate peripheral participation*. Cambridge university press.
40. Amanda Lenhart, Joseph Kahne, Ellen Middaugh, Alexandra Rankin Macgill, Chris Evans, and Jessica Vitak. 2008. Teens, Video Games, and Civics: Teens' Gaming Experiences Are Diverse and Include Significant Social Interaction and Civic Engagement. *Pew internet & American life project* (2008).
41. Jeffrey Lin. 2013. The science behind shaping player behavior in online games. In *Annual Game Developers Conference*.
42. Brendan Maher. 2016. Can a video game company tame toxic behaviour? *Nature* 531, 7596 (mar 2016), 568–571. DOI : <http://dx.doi.org/10.1038/531568a>
43. Yong Ming Kow, Timothy Young, and Katie Salen Tekinbas. 2014. Crafting the metagame: Connected learning in the Starcraft II community. (2014).
44. Virginia Molgaard and Richard Spoth. 2001. The Strengthening Families Program for young adolescents: overview and outcomes. *Residential Treatment for Children & Youth* 18, 3 (2001), 15–19. DOI : <http://dx.doi.org/10.1300/J007v18n03>
45. Matthew H. Rafalow and Katie Salen Tekinba. 2014. *Sackboy Planet : LittleBigPlanet 2 Players*. Technical Report. Digital Media and Learning Research Hub, Irvine, CA. 1–43 pages.
46. Kathryn E Ringland, Christine T Wolf, Lynn Dombrowski, and Gillian R Hayes. 2015. Making Safe: Community-Centered Practices in a Virtual World Dedicated to Children with Autism. In *CSCW '15*. ACM Press, New York, New York, USA, 1788–1800. DOI : <http://dx.doi.org/10.1145/2675133.2675216>
47. Travis L Ross and Andrew J Weaver. 2012. Shall We Play a Game? *Journal of Media Psychology* (2012).
48. Katie Salen. 2008. *The ecology of games: Connecting youth, games, and learning*. MIT press.
49. Joseph Seering, Robert E Kraut, and Laura Dabbish. 2017. Shaping Pro and Anti - Social Behavior on Twitch Through Moderation and Example - Setting. *Computer Support, Cooperative Work and Social Computing Proceedings* (2017), 111–125. DOI : <http://dx.doi.org/10.1145/2998181.2998277>

50. Kenneth B Shores, Yilin He, Kristina L Swanenburg, Robert Kraut, and John Riedl. 2014. The identification of deviance and its impact on retention in a multiplayer game. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. ACM, 1356–1365.
51. Myrna B Shure. 2001. I can problem solve (ICPS): An interpersonal cognitive problem solving program for children. *Residential Treatment for Children & Youth* 18, 3 (2001), 3–14.
52. Myrna B. Shure and George Spivack. 1980. Interpersonal problem solving as a mediator of behavioral adjustment in preschool and kindergarten children. *Journal of Applied Developmental Psychology* 1, 1 (dec 1980), 29–44. DOI : [http://dx.doi.org/10.1016/0193-3973\(80\)90060-X](http://dx.doi.org/10.1016/0193-3973(80)90060-X)
53. Myrna B. Shure and George Spivack. 1982. Interpersonal problem-solving in young children: A cognitive approach to prevention. *American Journal of Community Psychology* 10, 3 (jun 1982), 341–356. DOI : <http://dx.doi.org/10.1007/BF00896500>
54. Petr Slovák and Geraldine Fitzpatrick. 2015. Teaching and developing social and emotional skills with technology. *ACM Transactions on Computer-Human Interaction (TOCHI)* 22, 4 (2015), 19.
55. Petr Slovak, Chris Frauenberger, and Geraldine Fitzpatrick. 2017. Reflective Practicum: A Framework of Sensitising Concepts to Design for Transformative Reflection. In *CHI'17*. ACM, 2696–2707. DOI : <http://dx.doi.org/10.1145/3025453.3025516>
56. Jonas Heide Smith. 2004. Playing dirty understanding conflicts in multiplayer games. *The Association of Internet Researchers* September (2004), 19–22.
57. K. Squire. 2006. From content to context: Videogames as designed experience. *Educational Researcher* 35, November 2006 (2006), 19–29. DOI : <http://dx.doi.org/10.3102/0013189X035008019>
58. Constance Steinkuehler and Esra Alagoz. 2010. Out-of-School Virtual Worlds Based Programs : A Cross-Case Analysis Methods : Data Collection & Analysis. 2 (2010), 304–305.
59. Ca Steinkuehler, Sean Duncan, and David Simkins. 2007. Massively Multiplayer Online Games & Education: An Outline of Research. *Proceedings of the Eighth Conference of Computer Supported Collaborative Learning* 48, 1 (2007), 674–684.
60. Constance Steinkuehler, Kurt Squire, and Sasha a. Barab. 2012. *Games, Learning, and Society Learning and Meaning in the Digital Age*. Cambridge University Press. 490 pages.
61. Constance A Steinkuehler. 2002. Learning in Massively Multiplayer Online Games. *Proceedings of the 6th International Conference on Learning Sciences (ICLS '04)* 608 (2002), 521–528.
62. Tina L Taylor. 2009. *Play between worlds: Exploring online game culture*. Mit Press.
63. John A Velez, Chad Mahood, David R Ewoldsen, and Emily Moyer-Gusé. 2014. Ingroup versus outgroup conflict in the context of violent video game play: The effect of cooperation on increased helping and decreased aggression. *Communication Research* 41, 5 (2014), 607–626.
64. K. Weare and M. Nind. 2011. Mental health promotion and problem prevention in schools: what does the evidence say? *Health Promotion International* 26, S1 (nov 2011), i29–i69. DOI : <http://dx.doi.org/10.1093/heapro/dar075>
65. C Webster-Stratton and T Taylor. 2001. Nipping early risk factors in the bud: preventing substance abuse, delinquency, and violence in adolescence through interventions targeted at young children (0-8 years). *Prevention science* 2, 3 (2001), 165–192. DOI : <http://dx.doi.org/10.1023/A:1011510923900>
66. James V Wertsh and Peter Tulviste. 1990. Apprenticeship in thinking: Cognitive development in social context. *Science* 249, 4969 (1990), 684–686.