Available on: https://jsshrs.ep-journals.org/ E-ISSN: 3051-3561, P-ISSN: 3051-3553

Journal of Social Science and Human Research Studies

Volume : 02 Issue : 08 August-2026

# The Psychology of Keyboard Warriors Through David J. Pollay's Garbage Truck Theory

**Mohamad Murad Mahyudin Muheji (Ph.D.)**

Universiti Malaya Centre for Continuing Education (UMCCed), Universiti Malaya

**Published Date: August 15, 2026** **Article DOI: 10.65150/EP-jsshrs/V2E8/2026-06**

**ABSTRACT:** This conceptual article examines the psychology of *keyboard warriors* through David J. Pollay's *Garbage Truck Theory*. The article addresses digital aggression as a complex phenomenon shaped by emotional states, digital communication conditions, personality tendencies, cognitive processes and social interaction. The article draws on existing literature on *online disinhibition*, emotional factors, *Dark Tetrad* traits, *moral disengagement*, *emotional contagion* and cyberbullying. It extends Pollay's metaphor from everyday social interaction to digital communication. The analysis conceptualises digital aggression through three interconnected positions and one response dimension. The *garbage carrier* represents an individual who enters an interaction with accumulated negative experiences or emotions. The *garbage dumper* represents the transfer of negative emotional effects towards another user. The *garbage receiver* represents the individual exposed to such aggression. *Refusal* represents the recipient's capacity to avoid absorbing or reproducing the emotional negativity. This conceptual extension highlights the role of emotional regulation in interrupting hostile interaction. The article argues that *Garbage Truck Theory* offers a useful interpretive lens for connecting individual psychological experiences with the social dynamics of digital aggression. However, the theory remains a conceptual metaphor and does not establish causal relationships or imply that all *keyboard warriors* share similar psychological conditions. The article provides a theoretical basis for understanding digital aggression while identifying directions for future empirical investigation.

**KEYWORDS:** Digital Aggression, Garbage Truck Theory, Keyboard Warriors, Moral Disengagement, Online Disinhibition

## 1. INTRODUCTION

Social media has transformed the way people communicate and interact with one another. Digital platforms now allow individuals to express opinions, emotions and disagreements within seconds. This environment has also created new forms of aggressive communication. Among these is the growing presence of *keyboard warriors*, a term commonly associated with users who express hostility, provocation or personal attacks through digital platforms. The behaviour of *keyboard warriors* extends beyond ordinary disagreement. Some users respond to different opinions with insults, hostile remarks or deliberate provocation. Digital communication can make such behaviour easier because users do not always face the immediate social consequences of their words. Ooi et al. (2019) associate the behaviour of *keyboard warriors* with aggressive conduct in social media environments. Their work highlights the importance of examining the psychological conditions that may shape aggressive online interaction.

Psychological factors provide an important perspective for understanding this behaviour. Suler (2004) explains that online communication can reduce some of the social restraints normally present in face-to-face interaction. Anonymity, physical distance and the absence of direct interpersonal cues may encourage individuals to express thoughts and emotions more freely. These conditions do not explain every form of online aggression. They nevertheless provide an important context for understanding why some users behave differently online. Previous research also connects aggressive online behaviour with *trolling* and antisocial interaction. Cheng et al. (2015) found that antisocial behaviour can emerge within online discussion communities. Cheng et al. (2017) further demonstrated that negative experiences within online environments may influence later trolling behaviour. These findings suggest that aggressive digital behaviour develops through more than individual characteristics. The interaction between emotional states, personal tendencies and the digital environment also deserves attention.

David J. Pollay's *Garbage Truck Theory* offers a different way to approach this issue. Pollay (2010) describes people as carrying emotional "garbage" such as frustration, anger and disappointment. When this burden accumulates, some individuals look for opportunities to unload it onto others. Pollay's central message is that people do not have to accept this emotional burden from others.

> *"Many people are like garbage trucks. They run around full of garbage, full of frustration, full of anger, and full of disappointment. As their garbage piles up, they look for a place to dump it. And if you let them, they'll dump it on you. So when someone wants to dump on you, don't take it personally. Just smile, wave, wish them well, and move on. Believe me. You'll be happier."*
>
> Pollay, 2010

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

Pollay developed this metaphor in the context of everyday human interaction. This article extends its conceptual use to digital communication. The *keyboard warrior* provides a relevant context because online aggression may involve the transfer of negative emotions from one user to another. The metaphor therefore offers a useful starting point for examining the relationship between emotional burden, aggressive expression and the response of the recipient. The application of Pollay's theory does not suggest that all *keyboard warriors* carry unresolved anger or frustration. Online aggression can arise from different motives and psychological tendencies. Research on *Dark Tetrad* traits, *moral disengagement* and *trolling* further demonstrates the complexity of hostile online behaviour (Buckels et al., 2014; Bandura, 1999, 2002). Pollay's metaphor is therefore used in this article as an interpretive lens rather than as a psychological diagnosis.

This article analyses the psychology of *keyboard warriors* through David J. Pollay's *Garbage Truck Theory*. The discussion examines how emotional burdens may precede aggressive online expression and how digital environments can facilitate the transfer of negative emotions. It also considers the position of recipients and their capacity to refuse further involvement in the cycle of hostility. The article contributes to the existing discussion by extending Pollay's metaphor from everyday social interaction to digital communication. It conceptualises the interaction through the positions of the *garbage carrier*, the *garbage dumper* and the *garbage receiver*. It further introduces *refusal* as an important dimension because recipients can choose not to absorb or reproduce the emotional negativity directed towards them. Through this approach, the article brings together psychological and social explanations of online aggression within a single conceptual reading of *keyboard warrior* behaviour.

## 2. LITERATURE REVIEW

### 2.1 *Keyboard Warriors*, *Trolling* and Digital Aggression

The emergence of *keyboard warriors* is closely associated with changes in the way people communicate in digital spaces. Users can now express their views to a broad audience without facing the recipient directly. This environment has also created opportunities for aggressive, provocative and hostile communication. Ooi et al. (2019) use the term *keyboard warriors* in their discussion of *cyberbullying* through mobile social media. The phenomenon needs to be distinguished from *trolling* because the two terms do not carry the same meaning. Hardaker (2010) describes *trolling* through elements of aggression, deception, disruption and the ability to provoke a response. The study also distinguishes *trolling* from *flaming*, which can emerge as a response to threats or conflict in computer-mediated communication. This distinction places provocation at the centre of understanding *trolling* behaviour.

*Cyberbullying* has a more specific scope because it involves aggressive behaviour directed towards another individual through communication technologies. Kowalski et al. (2014) synthesised research on *cyberbullying* with particular attention to its predictors and consequences. Not every harsh comment or verbal attack therefore constitutes *cyberbullying*. Some incidents may occur as spontaneous reactions without developing into a broader pattern of harassment. Digital aggression can also take the form of hostile language, open confrontation and attacks directed at individuals. *Verbal aggression* uses language to hurt, insult or demean another person. *Flaming* refers to hostile exchanges that occur within digital communication. *Personal attacks* shift attention away from an issue and towards an individual's identity or perceived weaknesses. Hardaker (2010) places *flaming* within broader discussions of conflict and impoliteness in computer-mediated communication.

Patterns of interaction can also shape the development of aggressive behaviour. Cheng et al. (2015) examined antisocial behaviour across online discussion communities. They found that antisocial users were more likely to elicit responses from other users. The study also found that excessively harsh community feedback could intensify antisocial behaviour. Cheng et al. (2017) extended this explanation by showing that emotional states and discussion contexts can influence the tendency to engage in *trolling*. Their findings indicated that negative mood and exposure to *troll* posts increased the likelihood of subsequent *trolling*. Even ordinary users may engage in such behaviour when particular circumstances encourage it.

These findings are important because they challenge the view that digital aggression originates only from a particular group of users. Some individuals may respond because of their emotional state or the context of an interaction. Others may deliberately seek provocation or disruption. The phenomenon of *keyboard warriors* is therefore better understood as a broad behavioural category that encompasses different forms of aggression and different motives. In this article, *keyboard warriors* refer to users who employ digital communication to attack, provoke or express hostility towards other users. This definition focuses on observable behaviour without assuming that all perpetrators share the same psychological characteristics. Such an approach is important because Cheng et al. (2017) demonstrated that situational conditions and context can encourage ordinary users to engage in *trolling*. These conceptual distinctions provide a foundation for examining the psychology underlying digital aggression. *Trolling* emphasises provocation and disruption. *Cyberbullying* focuses on digital aggression directed at a target and its potential consequences. In this article, *keyboard warriors* are defined as users who employ digital spaces to express hostility through various forms of communication. The next section examines the psychological mechanisms that help explain why digital environments can alter how individuals regulate and express such behaviour.

### 2.2 Psychological Mechanisms of Online Aggression

The digital environment alters several social conditions that normally regulate face-to-face communication. Users do not need to face recipients when posting comments or responses. This distance can reduce the social pressure normally present in physical encounters. Suler (2004) explains this condition through the concept of *online disinhibition effect*. Anonymity represents one condition that may contribute to reduced social restraint. Users who conceal their identities may feel less bound by social norms governing communication. Physical distance also reduces the need to face recipients' immediate reactions. Suler (2004) associate anonymity and invisibility with changes in how users express themselves online. However, *online disinhibition effect* does not necessarily produce negative behaviour. Suler (2004) distinguishes between *benign disinhibition* and *toxic disinhibition*. The former may encourage openness and personal disclosure. The latter may encourage expressions that exceed ordinary social boundaries. These expressions can include anger, insults and hostility. This distinction indicates that digital environments do not determine behaviour in a fixed manner.

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

Recent evidence further supports the relevance of online disinhibition to harmful digital behaviour. Yang et al. (2025) examined the relationship between online disinhibition and cyberbullying. Their findings indicate that online disinhibition forms part of the psychological context surrounding harmful digital behaviour. This evidence strengthens the relevance of digital disinhibition when examining online aggression. Within digital aggression, *toxic disinhibition* may become more evident when users enter interactions with heightened emotional states. Anger may increase the tendency to respond harshly to opposing views. Frustration may reduce tolerance towards emotionally challenging stimuli. Existing hostility may also lead users to interpret other people's actions more negatively. Emotional states may also develop before users enter a particular discussion. Negative experiences in everyday life may influence their emotional state when they enter digital spaces. Particular stimuli may then trigger stronger reactions than would occur under ordinary conditions. Cheng et al. (2017) found that negative mood and exposure to *trolling* increased users' tendency to engage in *trolling*.

When emotional intensity rises, impulsive reactions may influence users' decisions to respond immediately. Individuals may post responses before considering how their words affect recipients. The speed of digital communication facilitates such reactions because users can produce responses within a short period. Emotional reactions may therefore precede more deliberate consideration. Difficulty in managing emotional intensity leads to the concept of *emotional dysregulation*. This condition refers to difficulties in managing emotional experiences and expressions effectively. Users may respond in ways that appear disproportionate to the stimuli they encounter. Such reactions may emerge through insults, provocation or attacks against other users. In some situations, users may also redirect emotions from their original source towards another target. This process relates to *displacement of affect* because individuals may express emotions towards someone who did not cause them. Digital spaces provide accessible targets through comments, responses and posts. A user who appears during a period of heightened emotion may therefore receive an emotional response originating from another experience.

These processes illustrate the interaction between digital conditions and individual psychological states. *Online disinhibition effect* may reduce certain forms of social restraint. Anonymity and physical distance may reduce the pressure associated with face-to-face interaction. Anger, frustration and hostility may create conditions for aggressive responses. Spontaneous reactions and *emotional dysregulation* may further affect the ability to control such responses. This perspective indicates that digital aggression does not emerge from a single factor. Environmental conditions provide opportunities for expression with weaker social restraint. Emotional states provide motivation for particular responses. Emotional displacement helps explain why another user may become a target despite not causing the original anger. The next section examines personality traits and *moral disengagement* as additional psychological dimensions.

**2.3 Personality, *Moral Disengagement* and Malevolent Online Behaviour**

Psychological factors in digital aggression extend beyond emotional states and environmental conditions. Certain personality traits may help explain differences in users' tendencies to engage in malevolent behaviour. The concept of the *Dark Tetrad* provides a useful basis for examining these tendencies through *narcissism*, *Machiavellianism*, *psychopathy* and *sadism*. Buckels et al. (2014) associated *trolling* with *Dark Tetrad* traits, with particular attention to *sadism*. Individuals with such tendencies may derive enjoyment from disturbing other users. This finding suggests that some forms of digital aggression involve a desire to elicit negative reactions from a target. Craker and March (2016) added another dimension through their discussion of *negative social potency*. This trait refers to a tendency to use social interaction to produce negative effects on others. This perspective highlights the roles of power and manipulation in *trolling* behaviour. *Machiavellianism* is relevant in this context because it involves the strategic use of social interaction to achieve particular goals.

Sest and March (2017) examined *psychopathy*, *sadism* and empathy in *cyber-trolling*. *Psychopathy* reflects a malevolent tendency that differs from the pleasure associated with observing another person's suffering. *Sadism* places greater emphasis on deriving satisfaction from another person's negative experience. Limited empathy may further reduce sensitivity to the emotional state of the target. March (2019) extended this understanding by examining psychopathy, sadism, empathy and the motivation to cause harm. The findings further support the role of malevolent traits in the tendency to engage in *Internet trolling*. At the same time, empathy provides an important counterbalance because the ability to understand another person's experience may reduce the motivation to cause harm. Hidalgo-Fuentes et al. (2025) synthesised broader evidence through a meta-analysis of *online trolling* and *Dark Tetrad* traits. Their analysis strengthened the association between the four personality traits and the tendency to engage in *trolling*. This evidence is important because it brings findings from multiple studies together and provides a broader view of personality-related tendencies.

Although *Dark Tetrad* traits offer an important explanation, personality does not determine behaviour on its own. An individual may possess certain tendencies while still choosing not to engage in aggression. Aggressive behaviour also depends on how individuals evaluate the acceptability of their actions and their consequences for others. Moral judgment therefore becomes important at this stage. Bandura (1999) describes *moral disengagement* as a cognitive process through which individuals separate their actions from their moral standards. Individuals may frame harmful behaviour as justified or minimise its negative effects on the target. They may also shift responsibility to circumstances or other parties. These processes weaken moral self-regulation when actions conflict with personal standards. Bandura (2002) further developed this explanation through cognitive mechanisms that allow individuals to maintain a sense of moral agency. A person may harm others without viewing themselves as someone who has violated their moral principles. Such mechanisms create room for rationalisation before or after aggressive behaviour occurs.

In digital communication, this rationalisation may appear through explanations that reduce feelings of guilt. A perpetrator may regard an insult as a joke or view an attack as a justified reaction. They may also blame the target by claiming that the target caused the conflict. Such judgments reduce the individual's sense of responsibility for the consequences of their actions. Taken together, these perspectives provide a more comprehensive understanding of malevolent behaviour online. *Dark Tetrad* traits help identify personality tendencies that may encourage aggressive behaviour. *Moral disengagement* addresses the cognitive processes that allow individuals to set aside moral considerations. The two perspectives therefore serve different functions in explaining the development of such behaviour. This distinction is important when examining *keyboard warriors*. Not every perpetrator possesses *Dark Tetrad* traits. Not every attack reflects a desire to harm another person. Some actions

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

may arise from emotional states or particular interactional circumstances. Personality traits and moral processes therefore add further layers to the explanation of online aggression.

This perspective strengthens the use of *Garbage Truck Theory* as a conceptual lens. Personality traits may help explain tendencies that encourage malevolent behaviour. *Moral disengagement* helps explain how perpetrators set aside consideration of the consequences of their actions. Together, these perspectives help clarify the conditions that may precede the transfer of emotional "garbage" to another user.

### 2.4 *Emotional Contagion*, Reinforcement and Victim Impact

Emotions do not necessarily remain within the individuals who experience them. Social interaction creates opportunities for emotional states to spread from one person to another. Hatfield et al. (1993) describe this process through the concept of *emotional contagion*. Their later work further develops this perspective by examining how individuals respond to the emotional expressions and behaviours of others (Hatfield et al., 1994). This concept provides a useful basis for understanding emotional transmission in digital communication. Digital spaces can accelerate such processes through comments, replies and shared content. Comments expressed through anger or hostility may trigger responses from other users. Recipients may respond in similar ways and extend the conflict to additional participants. Such interactions can make negative emotions part of the broader atmosphere of a discussion. Cheng et al. (2015) examined antisocial behaviour within online discussion communities. Their research indicates that negative behaviour is associated with patterns of interaction among users. Exposure to such behaviour may also influence the atmosphere of online communities. This suggests that digital aggression develops through social interaction rather than through individual tendencies alone.

Cheng et al. (2017) provide further insight into the development of *trolling* in digital environments. They found that negative mood and exposure to *trolling* increased the likelihood of engaging in similar behaviour. Users' experiences within a discussion can therefore influence subsequent responses. Users who did not initially engage in *trolling* may also adopt similar behaviour under particular conditions. Attention from other users adds another dimension to this process. *Likes*, *replies* and forms of *engagement* can increase the visibility of particular comments or posts. Such responses do not necessarily indicate agreement with the content. However, the attention provides social feedback to the perpetrator and can make the behaviour more noticeable. This condition helps explain the reinforcement of behaviour. A perpetrator may repeat an action when provocation produces an engaging response. A comment that generates numerous replies may provide an indication that the action has successfully produced a reaction. The conflict may then develop through a cycle of action, response and attention. This cycle can also affect recipients of digital attacks. Smith et al. (2008) examined the effects of *cyberbullying* on the well-being of young people. Their findings indicate that experiences of digital aggression can affect victims' psychological well-being. Online aggression therefore extends beyond the exchange of words within a digital space.

Kowalski et al. (2014) synthesised extensive evidence concerning the factors and consequences associated with *cyberbullying*. Their review linked victimisation experiences with various psychological difficulties. These effects place recipients at an important position within the analysis of digital aggression. Victims are therefore not merely targets at the end of an interaction. Hu et al. (2021) focused specifically on the relationship between *cyberbullying* victimisation and depressive symptoms among adolescents. Their meta-analysis highlights the importance of victimisation experiences when examining the psychological consequences of digital aggression. Negative online experiences may therefore contribute to substantial emotional distress among targeted individuals. Lee et al. (2026) examined the relationship between *cyberbullying victimization* and mental health symptoms through longitudinal evidence. Their approach considers psychological consequences across time. This evidence further reinforces the importance of examining recipients when considering the consequences of digital aggression.

The literature therefore points towards an interconnected process. Negative emotions may spread through interaction. User responses can provide attention to particular forms of behaviour. Such attention may reinforce the tendency to continue a conflict. Recipients may then experience psychological consequences from these interactions. This process gives particular meaning to the concept of *dumping garbage* within *Garbage Truck Theory*. A perpetrator transfers negative emotional effects to another user through digital interaction. The recipient then encounters the consequences of that transfer. The recipient's response may also influence whether the conflict ends or continues. This perspective connects the perpetrator's behaviour with the recipient's experience within an ongoing social process.

## 3. THE PSYCHOLOGY OF *KEYBOARD WARRIORS* THROUGH *GARBAGE TRUCK THEORY*

The literature reviewed above suggests that digital aggression emerges from an interaction between environmental conditions, emotional states and psychological factors. Pollay's (2010) *Garbage Truck Theory* provides a conceptual lens for examining this interaction. Pollay describes individuals as carrying emotional "garbage" and seeking opportunities to unload it onto others. This metaphor offers a way to interpret digital aggression as the transfer of negative emotions and psychological strain towards particular targets. This interpretation does not apply to all *keyboard warriors*. Some individuals may enter digital interactions with anger, frustration, disappointment or stress arising from previous experiences. Certain stimuli may then trigger stronger reactions. Comments that challenge personal views may also become a source of hostility. The framework therefore considers the possibility of emotional transfer without treating it as a universal characteristic of *keyboard warriors*. The concept of *carrying garbage* describes the condition that may precede the expression of negative emotions towards others. Individuals carry experiences and emotions from different areas of their lives. This emotional load may have no connection with the user who eventually becomes the target. A new interaction may nevertheless provide an opportunity to express unresolved feelings. Under certain circumstances, another user may therefore experience the effects of emotions that originated elsewhere.

The shift from *carrying garbage* to *dumping garbage* occurs when an individual directs emotional negativity towards another person. In digital spaces, this transfer may appear through insults, provocation, aggressive comments and personal attacks. *Online disinhibition* helps explain the environmental conditions that facilitate such expression. Suler (2004) argues that features of online communication can reduce some of the social restraints normally present in face-to-face interaction. A digital target may have no connection with the original source of the perpetrator's emotions. A user may become a target because of an opposing opinion or because a particular post attracts attention. The perpetrator also does not need to confront the target physically. Anonymity and psychological distance can reduce some of the social pressure normally associated with

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

face-to-face conflict. These conditions create an easier channel for expressing negative emotions through digital communication. Provocation then connects the perpetrator's behaviour with the target's response. A perpetrator may post challenging or disruptive comments to elicit a particular reaction. Golf-Papez and Veer (2017) emphasise the importance of user responses when examining the development of *trolling*. Their perspective suggests that provocative behaviour does not operate in isolation because the recipient's response also shapes the subsequent interaction.

The target's response can influence the direction of a conflict without guaranteeing a particular outcome. A recipient may respond with anger or defend their position. They may also choose not to engage with the provocation. Any response still becomes part of the interaction and may create an opportunity for further attacks. Shachaf and Hara (2010) illustrate how troll behaviour can disrupt interactions within digital communities and elicit responses from other users. Social attention adds another layer to this cycle. *Likes*, *replies* and *engagement* can increase the visibility of digital content under certain conditions. A large number of responses can also give attention to the perpetrator even when other users reject the content. In some situations, such attention may sustain interest in the conflict. Aggressive behaviour may then recur through continued interaction. The cycle does not involve only the perpetrator and the original target. Other users may enter a discussion after seeing comments or responses that have generated conflict. They may support, challenge or intensify the provocation. Negative emotions may then spread to additional participants through a chain of interaction. This pattern is consistent with the concept of *emotional contagion* described by Hatfield et al. (1993). The perpetrator may also remain engaged when the interaction generates attention or other forms of social response. Repeated involvement does not in itself demonstrate the presence of a particular psychological problem. This article treats such repetition only as part of an interactional pattern that may sustain aggression. Recurrent social responses can provide stimulation that reinforces particular behaviours.

Recipients, meanwhile, experience the consequences of the attacks directed towards them. Research on *cyberbullying* indicates that digital aggression can affect psychological well-being. Smith et al. (2008) identified psychological consequences among young people who experienced *cyberbullying*. Kowalski et al. (2014) likewise synthesised evidence concerning the factors and consequences associated with such experiences. Hu et al. (2021) provided further evidence of an association between victimisation and depressive symptoms. The effects on recipients do not necessarily appear in the same form for every individual. Such experiences may involve emotional distress and disruptions to psychological well-being. Lee et al. (2026) examined the relationship between *cyberbullying victimisation* and mental health symptoms using longitudinal evidence. This literature reinforces the importance of considering recipients as part of the analysis rather than treating them merely as the final targets of aggression. Within Pollay's framework, recipients retain a degree of choice in how they respond to an attack. They do not need to treat every provocation as something that requires a response. Actions such as *mute*, *block* and *report* can help manage exposure to problematic interactions. A calm response may also reduce opportunities for conflict to develop through reciprocal emotional reactions. These strategies give practical meaning to Pollay's principle of refusing to accept another person's "garbage".

Refusing *garbage* does not mean that recipients should ignore every form of aggression. Threats to personal safety or behaviour that violates platform rules may require appropriate action. In such circumstances, reporting the incident and preserving relevant evidence can form part of responsible interaction management. Pollay's principle is more appropriately applied to managing the acceptance and spread of negative emotions rather than to ignoring harmful behaviour. Taken together, the analysis reveals a chain of interaction. Individuals may enter digital spaces carrying experiences and emotional states from earlier situations. Particular conditions may encourage them to express those emotions through aggressive behaviour. Targets may respond or choose to end the interaction. Social attention may sustain the conflict while negative emotions may spread to other users. At this point, *Garbage Truck Theory* offers a conceptual extension from everyday interaction to digital communication. Pollay's metaphor allows the perpetrator to be understood as a carrier and dumper of emotional "garbage", while the recipient becomes a party who can influence whether the transfer continues to shape the interaction. Digital environments add anonymity, physical distance, rapid responses and social attention to this process. The theory does not establish that all aggression originates from emotional burdens or that every *keyboard warrior* deliberately seeks targets for emotional release. These limitations are important because they preserve the metaphor as an interpretive lens rather than a psychological diagnosis of perpetrators.

## 4. DISCUSSION

### 4.1 *Garbage Truck Theory* as a Lens for Digital Aggression

*Garbage Truck Theory* offers a conceptual approach to digital aggression without reducing the phenomenon to a single psychological cause. Pollay (2010) uses the metaphor of "*garbage*" to describe negative burdens that individuals carry before unloading them onto others. In the context of *keyboard warriors*, this metaphor provides a useful reference for examining emotional states, personality tendencies and social responses within one conceptual perspective. The value of the theory lies in distinguishing internal states from the behaviour that may follow. Anger, frustration and disappointment may exist before an individual enters an interaction. These states do not automatically produce aggression. Suler (2004) helps explain how digital conditions such as anonymity and psychological distance may reduce social restraint. Together, these conditions provide a context for understanding why some individuals may express hostility more readily online.

This perspective also avoids assuming that all *keyboard warriors* share the same psychological background. Cheng et al. (2017) found that emotional states and interactional experiences can influence the tendency to engage in *trolling*. Lopes and Yu (2017) examined different characteristics and motives among individuals who engage in *trolling*. Verbalyte et al. (2022) offered a more nuanced account by questioning the simple portrayal of trolls as individuals driven solely by psychopathic tendencies. This diversity requires Pollay's metaphor to be applied with caution. Trolling also needs to be understood within the broader culture of online interaction. Phillips (2015) examined the relationship between *online trolling* and mainstream digital culture. His work places *trolling* within a wider cultural environment rather than treating it as behaviour completely separate from ordinary online communication. This perspective further cautions against viewing every troll as a psychologically distinct type of user. Differences in motivation also create space for considering the *Dark Tetrad*. Buckels et al. (2014) placed particular emphasis on *sadism* in explaining the pleasure that some individuals may derive from *trolling*. Craker and March (2016) associated the *Dark Tetrad* with

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

*negative social potency* and *trolling* behaviour. Sest and March (2017) and March (2019) examined *psychopathy*, *sadism* and empathy in *cyber-trolling*. Hidalgo-Fuentes et al. (2025) later synthesised broader evidence through a meta-analysis of *Dark Tetrad* traits and *online trolling*.

Personality traits alone, however, cannot explain why an individual chooses to attack another person at a particular moment. *Moral disengagement* provides an additional layer of explanation. Bandura (1999, 2002) explains how individuals can rationalise behaviour that conflicts with their moral standards through various cognitive mechanisms. In digital communication, a perpetrator may regard an insult as a joke or blame the target for the conflict. Such rationalisation can weaken the moral restraints that might otherwise prevent continued aggression. Variations in motivation also matter because not all forms of *trolling* serve the same purpose. Mao et al. (2023) distinguish between *proactive trolling* and *reactive trolling*. This distinction indicates that similar behaviours may arise from different motivations. Pollay therefore does not need to account for every specific motive. His metaphor is more useful for describing the transfer of negative effects to another person after a particular motive produces an aggressive response. This perspective also extends the analysis beyond the perpetrator to the social interaction that develops after an attack. Hatfield et al. (1993) describe *emotional contagion* as the spread of emotional states through interaction. Cheng et al. (2015) likewise demonstrate that antisocial behaviour develops within the context of community interaction. Responses from other users can sustain a conflictual atmosphere and create opportunities for negative behaviour to recur. Within this context, Pollay's metaphor does not refer only to a transfer from one perpetrator to one target.

Instead, "*garbage*" may move through a chain of interactions when users respond to, join or intensify a conflict. This development makes recipients part of the dynamics of digital aggression. At the same time, recipients do not necessarily need to continue the cycle. Choosing not to engage with provocation gives practical meaning to Pollay's principle of allowing the "*garbage truck*" to pass. This is where the article extends the conceptual use of *Garbage Truck Theory*. Pollay discusses social situations in which individuals encounter others who unload negative burdens onto them. This article extends that interpretation to digital spaces with their own communication characteristics. Anonymity, physical distance, rapid responses and social attention can alter how such transfers occur. Pollay's theory therefore serves as a meeting point between individual emotional experiences and the dynamics of digital interaction. This contribution does not suggest that *Garbage Truck Theory* replaces psychological theories of digital aggression. *Online disinhibition* explains communication conditions that reduce social restraint. *Dark Tetrad* identifies particular personality tendencies. *Moral disengagement* addresses the rationalisation of behaviour. *Emotional contagion* helps explain the movement of emotions through interaction. Pollay provides a metaphor that brings these elements together through the idea of carrying and transferring negative experiences.

This position also clarifies the limits of the theory. Pollay did not develop an empirical model specifically for *keyboard warriors*. The metaphor of "*garbage*" cannot determine the actual cause of every attack. Not every perpetrator experiences emotional distress. Not every aggressive act represents an attempt to release frustration. Some individuals may act for entertainment, provocation, social attention or particular personality tendencies. This article therefore does not use the theory to diagnose perpetrators. The strength of *Garbage Truck Theory* ultimately lies in its interpretive function. The theory shifts attention from the question of who is at fault towards how negative burdens move through interaction. This approach includes perpetrators, targets and other users who enter the conflict. It also creates space for understanding how ending a response can interrupt the continuation of negative interaction. *Garbage Truck Theory* can therefore serve as a conceptual lens that connects individual psychology with the social dynamics of digital communication. The theory does not provide a single explanation for why someone becomes a *keyboard warrior*. Instead, it brings several psychological explanations into a broader account of carrying, unloading and receiving negative experiences. This extension gives Pollay's metaphor its theoretical value in understanding digital aggression.

### 4.2 The Psychological Position of the *Garbage Receiver*

The position of the *garbage receiver* adds an important dimension to the analysis of *Garbage Truck Theory*. Pollay (2010) does not focus only on individuals who unload their "garbage" onto others. His principle also recognises the choices available to those who encounter such behaviour. In the digital context, this position helps explain how users can respond to attacks without necessarily taking on the perpetrator's emotional burden. Receiving a digital attack and accepting another person's emotional burden are not the same. A person may encounter an insulting comment without accepting the negative judgment expressed in it. Conversely, an individual may continue thinking about the attack and allow the experience to affect their emotional state. This distinction matters because the presence of an attack does not automatically determine the recipient's psychological response. Research on *cyberbullying* indicates that experiences of digital aggression can affect psychological well-being. Hu et al. (2021) associated victimisation experiences with depressive symptoms among adolescents. Lee et al. (2026) further examined the relationship between victimisation and mental health symptoms through longitudinal evidence. These findings reinforce the importance of considering victims' experiences when analysing digital aggression.

Within the conceptual framework of this article, *emotional regulation* refers to an individual's capacity to manage emotional responses after encountering negative stimuli. This concept does not require recipients to deny feelings of anger, sadness or distress. Instead, individuals acknowledge these emotions without allowing them to determine their entire response. This approach distinguishes emotional regulation from emotional suppression. This distinction produces an important conceptual separation. Receiving an attack is not the same as accepting the perpetrator's emotional burden. Recipients cannot control every aggressive comment that appears in digital spaces. They can, however, exercise some control over their level of engagement after an attack occurs. They may respond impulsively, pause before reacting or end the interaction. Choosing to end an interaction gives practical meaning to the principle of *refusing the garbage*. *Mute*, *block* and *report* are not merely technical tools. They provide ways to manage exposure and establish boundaries within digital interactions. Section 3 considers these strategies as part of the process of interrupting aggression. Here, the same strategies are examined from the perspective of regulating emotional involvement.

Recipients must also distinguish emotional regulation from ignoring threats. Serious, repeated or threatening attacks require appropriate action. Users may preserve evidence, use reporting functions or seek assistance through available channels. Pollay's principle should not encourage victims to remain silent in the face of harmful behaviour. The position of the *garbage receiver* is therefore not passive. Recipients have a role in determining whether a negative interaction ends at a particular point or continues to develop. A controlled response may not change the

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

perpetrator's behaviour. It can nevertheless reduce the recipient's involvement in the cycle of hostility. In this context, *emotional regulation* becomes part of maintaining psychological boundaries. The distinction between receiving an attack and accepting an emotional burden adds conceptual value to this article. Pollay uses the metaphor to encourage individuals not to take another person's "garbage" personally. This article extends that idea to digital communication by distinguishing exposure to an attack from psychological acceptance of that attack. This distinction helps explain why recipients retain agency even when they cannot control the perpetrator's behaviour.

The *garbage receiver* is therefore not merely the final destination in a process of emotional transfer. Recipients occupy a psychological position that shapes how they manage the experience. They may recognise an attack without adopting the perpetrator's emotional state. They may also acknowledge the impact of an attack without continuing the hostility. In this sense, *refusing the garbage* represents a form of managing engagement rather than simply ignoring the perpetrator.

**4.3 Social and Legal Implications**

Digital aggression extends beyond disputes between individual users. Such behaviour can influence the broader atmosphere of interaction within digital communities. Cheng et al. (2015) found that antisocial behaviour within online discussion communities can develop through patterns of user interaction. This suggests that a single aggressive act may influence other users who later enter the discussion. Social consequences also depend on how users respond to aggressive behaviour. Hostile replies can prolong a conflict, while attention directed towards provocation may create opportunities for such behaviour to continue. Ooi et al. (2019) associated *keyboard warrior* behaviour with aggression in social media use. This evidence supports the need to view digital aggression as a social communication issue rather than merely a dispute between individuals. The implications for recipients also require attention. Experiences of digital attacks may create psychological distress, particularly when individuals face repeated attacks or extensive exposure. This perspective is consistent with the position of the *garbage receiver* discussed earlier. Recipients not only encounter aggressive content but must also manage their emotional responses and decide whether to continue or end the interaction. From a legal perspective, some forms of digital aggression may reach a level that carries legal consequences. The *Convention on Cybercrime*, also known as the *Budapest Convention*, provides an international framework for cooperation in addressing crimes involving information and communication technologies (Council of Europe, 2001). This framework becomes relevant when digital conduct extends across national jurisdictions.

The legal dimension, however, serves only as supporting context in this article. The primary focus remains on the psychological and social dimensions underlying *keyboard warrior* behaviour. The existence of an international legal framework simply indicates that certain forms of digital aggression may extend beyond ordinary interpersonal communication and require institutional intervention. These implications highlight the need to address digital aggression at several levels. Users need to regulate their emotional responses and establish appropriate interactional boundaries. Digital communities also need communication norms that discourage provocation and sustained hostility. Authorities have a role when particular forms of conduct reach a level that warrants legal intervention. This layered approach maintains *Garbage Truck Theory* as a psychological and social framework without turning the article into a legal analysis.

## 5. CONCLUSION

This article examined the psychology of *keyboard warriors* through Pollay's (2010) *Garbage Truck Theory*. The analysis suggests that the metaphor of "garbage" provides a useful way to understand the transfer of negative emotions in digital communication. Digital aggression does not occur as an isolated act. Instead, it involves an interplay between emotional states, digital conditions, personality tendencies, cognitive processes and social responses. The application of *Garbage Truck Theory* also indicates that digital aggression has no single psychological explanation. *Online disinhibition* helps explain communication conditions that facilitate aggressive expression. Emotional factors draw attention to anger, frustration and hostility. *Dark Tetrad* traits add a personality dimension to the analysis. *Moral disengagement* addresses the cognitive processes involved in rationalising aggressive behaviour. *Emotional contagion* extends the analysis to the spread of emotions through social interaction. The main conceptual contribution of this article lies in extending Pollay's metaphor to four positions within digital communication. The *garbage carrier* refers to an individual who enters an interaction carrying negative experiences and emotions. The *garbage dumper* refers to an individual who transfers those negative effects to another person. The *garbage receiver* refers to the person who encounters this emotional transfer. These three positions are complemented by *refusal*, which represents the recipient's choice not to accept or reproduce the "garbage".

*Refusal* adds an important dimension to the use of the theory. Recipients do not necessarily have to become the endpoint of emotional transfer. They retain choices regarding exposure, response and involvement in a conflict. *Emotional regulation* therefore becomes part of maintaining psychological boundaries. This approach distinguishes between experiencing an attack and accepting the perpetrator's emotional burden. This extension represents the main novelty of the article. The analysis does not merely use *Garbage Truck Theory* to portray *keyboard warriors* as individuals who unload negative emotions onto others. Instead, it develops the metaphor as a process involving a carrier, a dumper, a receiver and refusal. This framework connects individual psychological experiences with the dynamics of digital interaction without assuming that all perpetrators share the same motives or psychological states. The article nevertheless has several limitations. *Garbage Truck Theory* is a conceptual metaphor rather than a formal empirical theory of psychology. The metaphor of "garbage" therefore cannot determine the actual cause of every aggressive act. This article also does not establish that all *keyboard warriors* carry emotional burdens before attacking other users. Some forms of behaviour may arise from entertainment, provocation, social attention or particular personality tendencies.

The article does not establish causal relationships between psychological factors and digital aggression either. The analysis is conceptual and seeks to organise existing explanations through Pollay's perspective. This approach allows the theory to function as an interpretive lens without making empirical claims that the article cannot substantiate. Future research could empirically examine the positions of the *garbage carrier*, *garbage dumper*, *garbage receiver* and *refusal* within digital communication. Researchers could also examine the relationship between emotional transfer, *online disinhibition*, *Dark Tetrad* traits, *moral disengagement* and recipient responses. Empirical approaches could determine how far this conceptual structure reflects users' actual experiences across different digital environments. Overall, *Garbage Truck Theory* offers a simple yet valuable perspective for understanding the psychology of *keyboard warriors*. The theory does not explain every form of digital aggression and

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

does not replace more specific psychological theories. Its value lies in connecting the perpetrator's internal experiences with emotional transfer, recipient responses and the continuation of conflict. In digital environments where emotions can spread rapidly, recognising the "garbage", refusing to accept it and avoiding its further transmission represent important dimensions of managing human interaction.

**REFERENCES**


1) Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review, 3*(3), 193–209. https://doi.org/10.1207/S15327957PSPR0303_3
2) Bandura, A. (2002). Selective moral disengagement in the exercise of moral agency. *Journal of Moral Education, 31*(2), 101–119. https://doi.org/10.1080/0305724022014322
3) Buckels, E. E., Trapnell, P. D., & Paulhus, D. L. (2014). Trolls just want to have fun. *Personality and Individual Differences, 67*, 97–102. https://doi.org/10.1016/j.paid.2014.01.016
4) Cheng, J., Bernstein, M., Danescu-Niculescu-Mizil, C., & Leskovec, J. (2017). Anyone can become a troll: Causes of trolling behavior in online discussions. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 1217–1230. https://doi.org/10.1145/2998181.2998213
5) Cheng, J., Danescu-Niculescu-Mizil, C., & Leskovec, J. (2015). Antisocial behavior in online discussion communities. *Proceedings of the International AAAI Conference on Web and Social Media, 9*(1), 61–70. https://doi.org/10.1609/icwsm.v9i1.14583
6) Council of Europe. (2001). *Convention on cybercrime (Budapest Convention, ETS No. 185).* https://www.coe.int/en/web/cybercrime/the-budapest-convention
7) Craker, N., & March, E. (2016). The dark side of Facebook®: The Dark Tetrad, negative social potency, and trolling behaviours. *Personality and Individual Differences, 102*, 79–84. https://doi.org/10.1016/j.paid.2016.06.043
8) Golf-Papez, M., & Veer, E. (2017). Don't feed the trolling: Rethinking how online trolling is being defined and combated. *Journal of Marketing Management, 33*(15–16), 1336–1354. https://doi.org/10.1080/0267257X.2017.1383298
9) Hardaker, C. (2010). Trolling in asynchronous computer-mediated communication: From user discussions to academic definitions. *Journal of Politeness Research, 6*(2), 215–242. https://doi.org/10.1515/jplr.2010.011
10) Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). Emotional contagion. *Current Directions in Psychological Science, 2*(3), 96–99. https://doi.org/10.1111/1467-8721.ep10770953
11) Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1994). *Emotional contagion*. Cambridge University Press.
12) Hidalgo-Fuentes, S., González-Pérez, E., & Martínez-Álvarez, I. (2025). Relationship between online trolling and Dark Tetrad personality traits: A meta-analysis. *Psicothema, 37*(4), 29–41. https://doi.org/10.70478/psicothema.2025.37.26
13) Hu, Y., Bai, Y., Pan, Y., & Li, S. (2021). Cyberbullying victimization and depression among adolescents: A meta-analysis. *Psychiatry Research, 305*, 114198. https://doi.org/10.1016/j.psychres.2021.114198
14) Kowalski, R. M., Giumetti, G. W., Schroeder, A. N., & Lattanner, M. R. (2014). Bullying in the digital age: A critical review and meta-analysis of cyberbullying research among youth. *Psychological Bulletin, 140*(4), 1073–1137. https://doi.org/10.1037/a0035618
15) Lee, J., Choo, H., Zhang, Y., Cheung, H. S., Zhang, Q., & Ang, R. P. (2026). Cyberbullying victimization and mental health symptoms among children and adolescents: A meta-analysis of longitudinal studies. *Trauma, Violence, & Abuse, 27*(2), 391–406. https://doi.org/10.1177/15248380241313051
16) Lopes, B., & Yu, H. (2017). Who do you troll and why: An investigation into the nature of the online troll. *Personality and Individual Differences, 112*, 65–72. https://doi.org/10.1016/j.paid.2017.02.045
17) Mao, Y., Xu, T., & Kim, K. J. (2023). Motivations for proactive and reactive trolling on social media: Developing and validating a four-factor model. *Social Media + Society, 9*(4). https://doi.org/10.1177/20563051231203682
18) March, E. (2019). Psychopathy, sadism, empathy, and the motivation to cause harm: New evidence confirms malevolent nature of the Internet troll. *Personality and Individual Differences, 141*, 133–137. https://doi.org/10.1016/j.paid.2019.01.001
19) Ooi, K.-B., Lee, V.-H., Hew, J.-J., & Lin, B. (2019). Mobile social cyberbullying: Why are keyboard warriors raging? *Journal of Computer Information Systems, 61*(4), 371–382. https://doi.org/10.1080/08874417.2019.1679685
20) Phillips, W. (2015). *This is why we can't have nice things: Mapping the relationship between online trolling and mainstream culture*. MIT Press.
21) Pollay, D. J. (2010). *The law of the garbage truck: How to respond to people who dump on you, and how to stop dumping on others*. Sterling Publishing.
22) Sest, N., & March, E. (2017). Constructing the cyber-troll: Psychopathy, sadism, and empathy. *Personality and Individual Differences, 119*, 69–72. https://doi.org/10.1016/j.paid.2017.06.038
23) Shachaf, P., & Hara, N. (2010). Beyond vandalism: Wikipedia trolls. *Journal of Information Science, 36*(3), 357–370. https://doi.org/10.1177/0165551510365390
24) Smith, P. K., Mahdavi, J., Carvalho, M., Fisher, S., Russell, S., & Tippett, N. (2008). Cyberbullying: Its nature and impact in secondary school pupils. *Journal of Child Psychology and Psychiatry, 49*(4), 376–385. https://doi.org/10.1111/j.1469-7610.2007.01846.x
25) Suler, J. (2004). The online disinhibition effect. *CyberPsychology & Behavior, 7*(3), 321–326. https://doi.org/10.1089/1094931041291295
26) Verbalyte, M., Keitel, C., & Howard, K. (2022). Online trolls: Unaffectionate psychopaths or just lonely outcasts and angry partisans? *Politics and Governance, 10*(4), 396–407. https://doi.org/10.17645/pag.v10i4.5790
27) Yang, Y., Tao, S., & Du, W. (2025). Game avatar to keyboard warriors: The relationship between online disinhibition, internet gaming disorder and cyberbullying. *Acta Psychologica, 261*, 105753. https://doi.org/10.1016/j.actpsy.2025.105753


License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/