



Socrates' Three Filters and Bullying Prevention: An Exploratory Study of Ethical Speech in School Contexts

Kayro Panjaya¹, Juwaria Muqtadir²

^{1,2}Cendekia Harapan School, Bali, Indonesia

Abstract: Bullying remains a persistent educational problem that threatens students' well-being, classroom safety, and the moral climate of school life. Although many school-based anti-bullying strategies emphasize rules, supervision, counseling, and disciplinary responses, less attention has been given to the ethical quality of students' everyday speech. Previous studies have examined bullying through behavioral, psychological, social-emotional, and school-climate perspectives, but the use of philosophical tools to help students evaluate speech before it becomes harmful remains underexplored. This study examines Socrates' Three Filters of truth, goodness, and usefulness as a framework for strengthening ethical communication in bullying prevention. Using an exploratory qualitative-dominant design, the study combined a focused literature review with a small descriptive survey involving 10 upper-primary and lower-secondary students in an integrated basic education setting. The findings show that 90% of participants had witnessed or experienced bullying, 80% were unsure whether information shared in bullying-related comments was true, 50% were unsure whether they usually considered the necessity of their words before speaking, and 60% believed that applying the Three Filters could help reduce bullying. These findings suggest that harmful student communication is shaped not only by intentional aggression, but also by weak verification habits, limited reflection, and the normalization of unkind speech as honesty. The study contributes a conceptual-pedagogical framework that positions the Three Filters as a micro-ethical tool within broader school-based strategies. However, because the study involved a small, context-specific sample and did not test a formal intervention, its findings should be interpreted as exploratory rather than generalizable.

Keywords: Bullying prevention; Ethical communication; Respectful schools; School climate; Socrates' Three Filters

1. Introduction

Bullying remains one of the clearest indicators of relational harm in educational settings. It affects students' sense of safety, willingness to participate in class, and understanding of their place within the school community. Beyond immediate emotional distress, bullying is associated with reduced well-being, academic disruption, absenteeism, and damaged peer relationships (Chiracu et al., 2024; UNESCO, 2025). It should therefore be understood not merely as a disciplinary issue, but as a problem closely related to educational purpose, school climate, and character formation.

In contemporary research, bullying is commonly defined as repeated aggressive behavior involving an imbalance of power between perpetrator and victim (Olweus, 2013). It may take verbal, relational, social, physical, and increasingly digital forms. Although this definition is well established, the practical challenge in schools remains significant. Students often normalize harmful expressions as jokes, honesty, or ordinary peer interaction, while adults may notice overt conflict but overlook the cumulative effects of ridicule, exclusion, rumor-spreading, and dismissive comments. As a result, some forms of bullying persist in the space between what students experience as harmful and what schools formally recognize as reportable misconduct.

Global and national evidence shows that this problem remains substantial. UNESCO (2025) reports that school violence and bullying continue to undermine learning, partic-

Correspondence:

Name: Juwaria Muqtadir

Email: juwaria@cendekiaharapan.sch.id

Received: Apr 21, 2026;

Revised: Apr 23, 2026;

Accepted: May 16, 2026;

Published: May 30, 2026;



Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons

Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/>).

ipation, and educational inclusion across countries. Recent studies also link bullying and weak school climate to poorer academic, behavioral, and social outcomes (Chiracu et al., 2024; Kelly et al., 2024). In Indonesia, research similarly indicates that bullying is shaped not only by individual behavior, but also by interpersonal, institutional, and contextual factors. Kapile et al. (2023) note that bullying disrupts the learning environment and weakens students' confidence and social functioning, while Ramadhanti and Hidayat (2022) show that prevention requires not only reactive measures after incidents occur, but sustained guidance, supervision, and value formation. Other recent Indonesian studies further emphasize the roles of peer norms, school climate, and protective social conditions in shaping bullying perpetration, victimization, and student well-being (Borualogo et al., 2024, 2025; Zaneva et al., 2025).

In the Indonesian school context, verbal and relational bullying should be understood in relation to local peer interaction, school climate, and social norms. A systematic review on bullying in Indonesia emphasizes that bullying cannot be separated from the local social and cultural practices in which it occurs (Tas'adi & Fanany, 2023). Recent empirical studies also show that Indonesian children may experience bullying through being called unkind names, being excluded by classmates, peer conflict, and negative school climate (Borualogo et al., 2024). In addition, evidence from the ROOTS-Indonesia intervention indicates that when negative interpersonal behaviors are perceived as normative by students, bullying perpetration and victimization are more likely to occur (Zaneva et al., 2025). Therefore, harmful speech may not always be immediately recognized as bullying, especially when it is framed as joking, teasing, honesty, or ordinary peer interaction. This suggests that bullying prevention in Indonesian schools requires not only rules against aggressive behavior, but also educational tools that help students evaluate the moral quality and social consequences of their everyday speech.

Many anti-bullying efforts therefore focus on structural and behavioral strategies such as supervision, school rules, sanctions, counseling, classroom management, and whole-school programs. These remain essential. However, such approaches do not always foreground a basic educational question: how should students learn to evaluate what they are about to say to another person? This question matters because verbal and relational bullying often begins in ordinary speech. A degrading nickname, humiliating remark, rumor, or public insult is not only a behavioral incident. It is also a failure of ethical judgment in communication.

This concern creates space for a philosophical and pedagogical contribution. Although Socrates did not formulate a modern anti-bullying theory, his ethical orientation toward examined speech, reflective judgment, and moral responsibility remains relevant to contemporary education. In educational terms, Socrates' Three Filters can be interpreted as a practical moral reasoning tool that encourages students to pause and examine whether a statement is true, whether it is good or respectful, and whether it is useful or constructive before it is spoken. This interpretation is consistent with the Socratic emphasis on examining claims, motives, and consequences (Brickhouse & Smith, 1990; Vlastos, 1991). It also aligns with modern educational concerns about critical thinking, character formation, social-emotional learning, and responsible communication (Mammen & Paulus, 2023; Meyer, 2024). The filter of truth strengthens students' ability to question rumor, exaggeration, and distortion; the filter of goodness directs attention to dignity, empathy, and respect; and the filter of usefulness asks whether speech contributes constructively to the situation rather than merely producing embarrassment, exclusion, or harm.

The theoretical value of the Three Filters therefore lies not in treating them as a complete anti-bullying theory, but in using them as a micro-ethical framework for everyday communication. This framework supports students' moral reasoning by transforming abstract values into concrete questions that can be applied before speaking: Is this true? Is this good? Is this useful? Moral reasoning develops through social interaction and can be strengthened through educational practices that encourage students to justify, examine, and communicate ethical judgments (Mammen & Paulus, 2023; Meyer, 2024).

However, the Three Filters also need to be connected with contemporary educational theories. Social learning theory suggests that students' speech behavior is shaped by peer reinforcement, imitation, and group approval (Bandura, 1977), while social-emotional learning emphasizes self-management, social awareness, relationship skills, and responsible decision-making as important conditions for constructive peer interaction (Fredrick et al., 2022). Thus, the Three Filters are best understood as a reflective communication tool that becomes educationally meaningful when supported by empathy, teacher guidance, classroom dialogue, and a positive school climate.

This framework is especially relevant to verbal and relational bullying, because harmful peer communication often appears through name-calling, rumors, personal insults, humiliating comments, and social labeling that can produce negative emotional reactions among students (Kapitanoff & Pandey, 2024). A degrading rumor fails truth. A humiliating insult fails goodness. A mocking public comment fails usefulness because its purpose is not to improve the situation, but to embarrass, dominate, or entertain at another student's expense. In this sense, the Three Filters offer a compact moral language that may help students pause before speaking and reflect on the consequences of their words.

However, philosophical clarity alone does not produce behavioral change. Students may understand what respectful speech looks like and still violate it under peer pressure, anger, retaliation, or the desire for status. This possibility is also consistent with research on moral disengagement, which explains how young people may justify harmful behavior even when it conflicts with moral standards (Luo & Bussey, 2023). Bullying is shaped not only by individual choice, but also by peer norms, emotional competencies, school connectedness, classroom climate, adult responses, and broader patterns of social learning (Bandura, 1977; Fredrick et al., 2022; Gaffney et al., 2021b). If harmful speech is rewarded with laughter, attention, or group approval, ethical reflection may be displaced by social reinforcement. This helps explain why a purely individual filter-based approach is too narrow to address the wider dynamics of bullying.

For this reason, the educational value of Socrates' Three Filters depends on contextual support. Empathy is one important complement, because it helps students consider not only whether a statement is accurate, but also how it affects another person. Recent literature links empathy to defending behavior, prosocial responses, and lower tolerance of bullying (Hikmat et al., 2024). Experimental work also suggests that teacher messages combining condemnation of bullying with empathy-raising elements are more likely to encourage intentions to stop bullying (Johander et al., 2022). Relatedly, social-emotional learning research indicates that self-management, relationship skills, and positive school climate are associated with lower bullying victimization and more constructive peer interaction (Fredrick et al., 2022).

School climate and whole-school approaches are equally important. Positive school climate is associated with belonging, safety, prosocial relationships, and reduced aggression (La Salle-Finley et al., 2024). Conversely, environments marked by weak connectedness, inconsistent norms, poor teacher-student relationships, or tolerance of ridicule create conditions in which bullying is more likely to persist. Reviews of school-based anti-bullying programs show that whole-school interventions can reduce both perpetration and victimization, although their effectiveness depends heavily on implementation and school context (Gaffney et al., 2021a, 2021b; Pearce et al., 2024). Restorative approaches also offer a useful complement because they strengthen accountability, dialogue, relationship repair, and peaceful conflict resolution after harm occurs (Lodi et al., 2021; Pérez-Jorge et al., 2023).

The research gap addressed in this article lies in the limited integration between anti-bullying research and philosophical frameworks for ethical speech. Existing anti-bullying studies have strongly explained bullying through behavioral, psychological, social-emotional, peer-norm, and school-climate perspectives. These studies have also shown the importance of empathy, teacher intervention, restorative practice, and whole-school strategies in reducing bullying. However, they have not sufficiently ex-

plained how students can be guided to evaluate the ethical quality of everyday speech before it becomes harmful. This limitation is important because verbal and relational bullying often begins not as a formal disciplinary incident, but as an ordinary comment, joke, rumor, nickname, or expression of “honesty” that gradually harms another student’s dignity. Conversely, Socrates’ Three Filters are often presented as personal moral wisdom, but they are rarely examined as a pedagogical framework for bullying prevention in contemporary school contexts. This creates a practical and theoretical problem: anti-bullying programs may regulate harmful behavior after it occurs, but students may still lack a simple reflective language for deciding whether a comment should be spoken in the first place.

Based on this problem, the study is guided by three research questions. First, how can Socrates’ Three Filters be interpreted as a framework for ethical communication in bullying prevention? Second, how do students perceive the relevance of truth, goodness, and usefulness in relation to harmful speech and bullying? Third, what school-based strategies are needed to make the Three Filters educationally effective rather than merely individual moral advice?

Therefore, this article aims to examine the educational relevance and limitations of Socrates’ Three Filters for bullying prevention in schools. It argues that the Three Filters should not be treated as a stand-alone solution, but as a micro-ethical framework that can support reflective communication when integrated with empathy-building, social-emotional learning, teacher-guided responses, restorative dialogue, and whole-school expectations for respectful interaction. The contribution of this study lies in connecting a classical ethical framework with contemporary school-based bullying prevention and in proposing ethical speech as an important but underdeveloped dimension of anti-bullying education.

2. Materials and Methods

2.1. Research Design

This study employed an exploratory qualitative-dominant design supported by a small descriptive survey. The design was selected because the study did not aim to test the statistical effectiveness of an intervention, but to explore how Socrates’ Three Filters may be interpreted and applied as a framework for ethical communication in bullying prevention. The qualitative component was dominant because the main purpose of the study was conceptual and interpretive, namely to connect the philosophical framework of truth, goodness, and usefulness with contemporary school-based bullying prevention. The descriptive survey served a supportive function by providing context-specific student perceptions of bullying, harmful speech, and reflective communication.

The study integrated two sources of evidence: a focused literature review and student response data. The literature review provided theoretical and empirical grounding on bullying, empathy, peer norms, school climate, social-emotional learning, restorative practice, and whole-school prevention strategies. The student survey provided exploratory insight into how learners perceive bullying-related speech and the possible relevance of the Three Filters in everyday school communication. Therefore, the study should not be understood as a causal intervention study, but as an exploratory conceptual-empirical study that uses student perceptions to enrich the interpretation of a pedagogical framework.

2.2. Research Setting and Participants

The study was conducted at Cendekia Harapan School, an integrated basic education institution located in Bali, Indonesia. This school context was selected because it represents an educational environment in which upper-primary and lower-secondary students interact regularly in classroom, peer-group, and informal school settings. These interactional spaces are relevant to the study because verbal teasing, nicknames, rumors, exclusion, and harmful comments often occur in everyday peer communication.

Although the school setting is identified, no individual participant names or personally identifiable student information are reported in this article.

The participants consisted of 10 students from upper-primary and lower-secondary levels. This age group was considered relevant because students at this level are developing peer identity, moral reasoning, social awareness, and communication habits. The small number of participants was appropriate for the exploratory purpose of the study, but it was not intended to represent the wider student population statistically.

2.3. Sampling Technique and Eligibility Criteria

Participants were selected using purposive sampling. This technique was used because the study required participants who were familiar with everyday peer interaction in school and able to reflect on bullying-related communication. The selection was not intended to produce statistical representativeness, but to obtain exploratory insight from students whose school experiences were relevant to the research focus.

The inclusion criteria were: students enrolled in upper-primary or lower-secondary levels, students who regularly participated in classroom and peer-group interactions, students who were able to understand the questionnaire items, and students who were willing to participate voluntarily. The exclusion criteria were: students who were absent during data collection, students who did not provide assent, students whose parents or guardians did not provide permission where required, and students who were unable or unwilling to complete the questionnaire.

2.4. Research Instrument

Data were collected using a structured questionnaire consisting of closed-ended and reflective items. The questionnaire covered four main areas: students' awareness and experience of bullying, students' assessment of the truthfulness of harmful comments, students' views on kindness and usefulness in communication, and students' reflection on whether respectful language might help reduce bullying. The instrument was designed to capture students' perceptions rather than to diagnose bullying behavior clinically or measure intervention outcomes.

The questionnaire was developed based on the conceptual structure of Socrates' Three Filters and relevant literature on bullying, ethical communication, empathy, and school climate. To support content relevance, the items were reviewed by educators familiar with student communication and bullying prevention. Because this was an exploratory study with a small sample, the questionnaire was not treated as a standardized psychometric scale. Therefore, statistical validity testing and internal consistency testing were not conducted. Instead, procedural reliability was supported through consistent wording, uniform administration, and alignment between questionnaire items and the conceptual dimensions of truth, goodness, and usefulness.

2.5. Data Collection Procedure and Research Ethics

Data collection was conducted through a classroom-friendly questionnaire procedure. Before completing the questionnaire, students were informed about the purpose of the study in age-appropriate language. They were also informed that participation was voluntary, that they could choose not to answer any item, and that their responses would be used only for research purposes. No identifying personal information was reported in the article.

Ethical considerations included voluntary participation, student assent, confidentiality of responses, and sensitivity to the topic of bullying. Where required, permission was obtained from the school and/or parents or guardians before data collection. Because the topic may involve uncomfortable experiences, the questionnaire avoided asking students to describe specific traumatic incidents in detail. The study focused on perceptions of communication and bullying prevention rather than on identifying individual perpetrators or victims.

2.6. Data Analysis

Closed-ended survey responses were analyzed descriptively using simple percentages to identify response patterns. Because the sample size was small, no inferential statistical analysis was conducted. The percentages were used only to describe tendencies within the participant group, not to make statistical generalizations.

Reflective responses were analyzed through a simple thematic coding process. First, responses were read repeatedly to identify recurring meanings. Second, initial codes were assigned to statements related to truth, kindness, usefulness, self-control, empathy, peer influence, and school response. Third, similar codes were grouped into broader themes. Fourth, these themes were compared with the findings from the focused literature review to identify points of convergence, tension, and limitation.

The coding process was interpretive rather than statistical. Its purpose was to clarify how students understood harmful speech and how their reflections related to the Three Filters. To reduce overinterpretation, the analysis was limited to recurring patterns that were directly supported by the questionnaire responses and interpreted cautiously in light of the small sample.

2.7. Literature Selection and Methodological Integration

The literature review was focused rather than fully systematic. It prioritized peer-reviewed journal articles, academic books, and major institutional reports related to bullying prevention, ethical communication, moral reasoning, empathy, school climate, social-emotional learning, restorative practice, and Indonesian educational contexts. Foundational works were included where necessary to frame Socratic ethics, social learning, and bullying theory.

Literature was selected based on four criteria: relevance to bullying prevention or harmful peer communication, relevance to ethical speech or moral reflection, relevance to school-based strategies, and relevance to Indonesian or comparable educational contexts where available. The purpose of the review was not to provide an exhaustive systematic review, but to construct a theoretically grounded interpretive framework for understanding the educational relevance of Socrates' Three Filters.

Methodological integration occurred at the interpretation stage. The descriptive survey results were used to identify student response patterns, while the literature review was used to interpret why these patterns matter for bullying prevention and how the Three Filters can be integrated with broader school-based strategies. In this way, the study combined context-specific student perceptions with established literature, while avoiding claims of statistical representativeness or causal effectiveness.

3. Results and Discussion

3.1. Results

The results are presented descriptively because the study involved a small exploratory sample of 10 students. Therefore, each percentage should be read as the number of students within this participant group rather than as a statistically generalizable estimate. For example, 90% represents 9 out of 10 students, 80% represents 8 out of 10 students, and 60% represents 6 out of 10 students. The analysis focuses on identifying response patterns related to bullying experience, truth-verification, kindness, usefulness, and students' perceptions of Socrates' Three Filters as a reflective communication tool.

a. Student Awareness and Experience of Bullying

Most participants reported direct or indirect exposure to bullying. Nine out of ten students stated that they had either witnessed or experienced bullying. In addition, four out of ten students were unsure whether they had ever bullied someone themselves. These response patterns are presented in Table 1.

Table 1. Student Awareness and Experience of Bullying

Indicator	Percentage
Had witnessed or experienced bullying	90%
Were unsure whether they had ever bullied someone	40%

These findings indicate that bullying was not a distant or abstract issue for the participants. The fact that 9 out of 10 students had witnessed or experienced bullying suggests that harmful peer interaction was part of their recognizable school experience. At the same time, the finding that 4 out of 10 students were unsure whether they had ever bullied someone themselves is analytically important. It suggests that students may not always recognize their own behavior as bullying, especially when harmful speech is framed as joking, teasing, honesty, or ordinary peer interaction. This uncertainty supports the need for a reflective framework that helps students evaluate their own communication before it harms others.

b. Student Responses Related to the Filter of Truth

Responses related to the filter of truth indicated considerable uncertainty. Six out of ten students reported that they had questioned negative comments made by bullies, and the same proportion said they had at times heard that bullying-related comments were untrue. At the same time, eight out of ten students were unsure whether the information shared by bullies was true. These findings are summarized in Table 2.

Table 2. Student Responses Related to the Filter of Truth

Item	Percentage
Questioned negative comments made by bullies	60%
Occasionally heard that bullying comments were untrue	60%
Were unsure whether the information shared by bullies was true	80%

The most notable finding in this section is that 8 out of 10 students were unsure whether information shared by bullies was true. This is significant because verbal and relational bullying often depends on unverified claims, rumors, exaggeration, or social labeling. When students are unsure about the truthfulness of harmful comments but still allow such comments to circulate, misinformation can become a tool of humiliation or exclusion. The truth filter is therefore relevant not only as a philosophical idea, but also as a practical educational prompt that encourages students to question whether a statement is factual before repeating or accepting it.

c. Student Responses Related to the Filter of Goodness

Responses related to the filter of goodness showed mixed views on kind and harmful speech. Some students recognized that hurtful language was ethically problematic, while others still considered unkind speech acceptable in certain situations. Four out of ten students believed that not all comments need to be kind because criticism is sometimes necessary. Half of the participants had heard bullying justified as “just being honest,” and half also agreed that bullies were aware that their words were unkind. Table 3 summarizes these responses.

Table 3. Student Responses Related to the Filter of Goodness

Item	Percentage
Believed not all comments need to be kind because criticism is sometimes necessary	40%
Had heard bullying justified as “just being honest”	50%
Agreed that bullies were aware their words were unkind	50%

The responses related to goodness reveal an important tension between honesty and kindness. Four out of 10 students believed that not all comments need to be kind because criticism is sometimes necessary, while 5 out of 10 had heard bullying justified as “just being honest.” This pattern suggests that students may struggle to distinguish constructive criticism from harmful speech. In bullying contexts, the claim of honesty can become

a moral justification for humiliating or degrading others. Therefore, the goodness filter is important because it does not reject truth-telling, but asks whether truth is communicated in a way that respects the dignity of others.

d. Student Responses Related to the Filter of Usefulness

Responses related to the filter of usefulness suggested that reflective consideration before speaking was not yet firmly established. Half of the students were unsure whether they usually asked whether a conversation or action was necessary before speaking. Seven out of ten were also unsure whether they would intervene by asking whether a conversation was useful, good, or truthful. These results are shown in Table 4.

Table 4. Student Responses Related to the Filter of Usefulness

Item	Percentage
Were unsure whether they usually ask if a conversation or action is necessary before speaking	50%
Were unsure whether they step in to question whether a conversation is useful, good, or truthful	70%

The responses related to usefulness suggest that purposeful reflection before speaking was not yet a stable habit among the participants. Five out of 10 students were unsure whether they usually considered whether a conversation or action was necessary before speaking, and 7 out of 10 were unsure whether they would intervene by asking whether a conversation was useful, good, or truthful. This finding is important because bystander silence may not only result from fear, but also from uncertainty about how to evaluate or interrupt harmful communication. The usefulness filter can therefore help students consider whether a comment contributes to problem-solving, relationship repair, or mutual understanding, rather than merely producing embarrassment, laughter, or exclusion.

e. Cross-Filter Pattern: Truth, Goodness, and Usefulness

When the three filters are read together, the findings suggest a possible descriptive pattern of weak reflective communication. The highest uncertainty appeared in the truth filter, where 8 out of 10 students were unsure whether bullying-related information was true. This uncertainty was accompanied by uncertainty in the usefulness filter, where 5 out of 10 students were unsure whether they usually considered the necessity of their words before speaking and 7 out of 10 were unsure whether they would intervene in conversations lacking truth, goodness, or usefulness. Although the data do not allow a statistical correlation between weakness in truth and weakness in usefulness, the pattern suggests that students who lack habits of verification may also lack a stable reflective language for deciding whether speech serves a constructive purpose.

This cross-filter pattern is educationally important. If students do not question whether a statement is true, they may be more likely to repeat rumors, labels, or negative comments. If they also do not ask whether the statement is useful, they may fail to consider whether the comment helps resolve a problem or merely harms another student. Therefore, the Three Filters should be understood as interconnected rather than separate questions. Truth helps prevent misinformation, goodness protects dignity, and usefulness directs speech toward constructive purposes.

f. Self-Reflection and Suggested Ways to Reduce Harmful Speech

Half of the students reported that they often reflected on whether their words were true, kind, and useful before speaking. When asked how to make communication more thoughtful, students selected several strategies, with the most common being educating themselves about the consequences of bullying, learning more about the impact of words, and reflecting before speaking. The distribution of these preferences is presented in Table 5.

Table 5. Student Preferences for More Thoughtful Communication

Preferred strategy	Percentage
Educating themselves on the consequences of bullying	70%
Learning more about the impact of words	60%
Reflecting before speaking	50%
Encouraging others to use kind language	30%

Students’ preferred strategies indicate that they did not view bullying prevention only as a matter of punishment or external control. Seven out of 10 students selected education about the consequences of bullying, 6 out of 10 selected learning more about the impact of words, and 5 out of 10 selected reflection before speaking. These responses suggest that students may be receptive to preventive approaches that combine knowledge, empathy, and self-reflection. However, only 3 out of 10 selected encouraging others to use kind language, which may indicate that students are less confident in influencing peers directly. This supports the need for teacher-guided and school-supported communication practices rather than relying only on individual student initiative.

g. Student Views on the Three Filters as a Preventive Tool

Most participants viewed Socrates’ Three Filters positively as a possible preventive tool in school. Six out of ten students (60%) believed that applying the Three Filters could help reduce bullying at school.

This finding should be interpreted cautiously. It does not prove that the Three Filters reduce bullying, but it indicates that most participants perceived the framework as understandable and potentially useful. In an exploratory study, this perception is important because a school-based ethical communication framework must first be meaningful to students before it can be developed into a structured intervention. The finding also suggests that the Three Filters may function as a simple language for classroom reflection, especially when supported by empathy-building, teacher guidance, and consistent school norms.

3.2 Discussion

a. Interpreting the Student Findings Cautiously

The discussion interprets the survey findings in relation to the focused literature review. Because the study involved only 10 students and used an exploratory design, the findings should not be interpreted as statistically generalizable evidence or as proof that Socrates’ Three Filters reduce bullying. Instead, the data provide context-specific insight into how a small group of students perceived bullying-related speech, uncertainty about truth, the tension between honesty and kindness, and the possible usefulness of reflective communication. The literature is then used to explain why these patterns matter and how the Three Filters may be integrated with broader school-based bullying prevention strategies.

The student data show three main descriptive patterns. First, bullying was a recognizable issue for most participants, as 9 out of 10 students reported that they had witnessed or experienced bullying. This finding suggests that bullying was not an abstract or distant issue for the participants, but part of their recognizable school experience. This pattern is consistent with broader evidence showing that bullying remains a persistent educational concern because it affects students’ sense of safety, participation, well-being, and learning conditions (UNESCO, 2025). It also aligns with Indonesian studies showing that bullying continues to influence students’ confidence, social functioning, school adjustment, and subjective well-being (Borualogo et al., 2024, 2025; Kapile et al., 2023).

Second, the findings indicate uncertainty in students’ evaluation of harmful speech, especially in relation to truth and usefulness. Eight out of 10 students were unsure whether information shared in bullying-related comments was true, while 5 out of 10 were unsure whether they usually considered the necessity of their words before speaking. This descriptive pattern is educationally important because verbal and relational

bullying often depends on rumors, exaggeration, social labeling, and repeated claims that may not be verified. A rumor does not need to be true in order to harm another student; it only needs to circulate within a peer group. Therefore, students' uncertainty about truth suggests the need for a simple reflective tool that can help them pause before repeating or accepting harmful claims.

Third, the findings show moderate openness toward Socrates' Three Filters, as 6 out of 10 students believed that applying the framework could help reduce bullying at school. This finding should be interpreted cautiously. It does not demonstrate the effectiveness of the Three Filters as an intervention, but it suggests that the framework may be understandable and meaningful to students as a language for reflection. In an exploratory study, this perception matters because any school-based ethical communication framework must first be accessible to students before it can be developed into a more structured pedagogical intervention.

Taken together, the survey findings support only a limited but important claim: the Three Filters appear relevant as a reflective language for discussing harmful speech among the participants. They do not demonstrate that the Three Filters can independently prevent bullying, nor do they establish a causal relationship between ethical reflection and behavioral change. Rather, they identify a pedagogical problem: students may encounter harmful speech without consistently verifying whether it is true, considering whether it is respectful, or asking whether it serves a constructive purpose.

b. Truth, Goodness, and Usefulness as Interconnected Dimensions of Ethical Speech

The findings suggest that the three filters should be understood as interconnected rather than separate moral questions. The truth filter addresses students' ability to evaluate whether a statement is factual before it is repeated. This is especially relevant because bullying-related communication often involves rumor, exaggeration, distortion, and social labeling (Kapitanoff & Pandey, 2024). When students do not question the truthfulness of harmful statements, misinformation can become a tool of humiliation and exclusion. In this sense, the truth filter supports not only factual accuracy, but also communicative responsibility.

The goodness filter addresses the ethical quality of speech. The finding that some students had heard bullying justified as "just being honest" indicates an important tension between honesty and kindness. Students may believe that truthful comments are automatically acceptable, even when those comments are delivered in a humiliating or degrading manner. However, truthful criticism is not the same as harmful speech. Educationally, the filter of goodness asks whether speech protects the dignity of others, whether it expresses disagreement respectfully, and whether it avoids unnecessary injury. This interpretation is consistent with the broader aim of character education, which does not merely teach students what is true, but also how truth should be communicated in socially responsible ways (Mammen & Paulus, 2023; Meyer, 2024).

The usefulness filter addresses the purpose of speech. The finding that 7 out of 10 students were unsure whether they would intervene by asking whether a conversation was useful, good, or truthful suggests that students may lack a stable language for interrupting harmful peer communication. Bystander silence may not only result from fear, but also from uncertainty about how to evaluate or challenge harmful speech. The usefulness filter therefore helps students consider whether a comment contributes to problem-solving, relationship repair, mutual understanding, or protection of others, rather than merely producing laughter, embarrassment, exclusion, or peer approval.

Although the data do not allow a statistical correlation between weakness in truth and weakness in usefulness, the pattern suggests a possible descriptive connection. Students who are unsure about the truthfulness of bullying-related information may also be unsure about whether such communication serves a constructive purpose. This does not prove a relationship between the two filters, but it indicates that verification and purposeful communication may need to be taught together. Truth helps prevent the spread

of misinformation, goodness protects dignity, and usefulness directs speech toward constructive outcomes.

c. Literature-Based Interpretation: Why the Three Filters Cannot Stand Alone

The literature helps explain why a reflective speech framework must be supported by broader educational strategies. Bullying is not only a matter of individual speech choice. It is shaped by peer norms, social reinforcement, school climate, adult responses, empathy, and the broader relational culture of the school. From a social learning perspective, students' behavior is influenced by imitation, peer approval, attention, laughter, and group reinforcement (Bandura, 1977). This means that even when students understand that a statement is unkind or unnecessary, they may still participate in harmful communication if such behavior gives them social status or acceptance. This dynamic can also be understood through moral disengagement, where students may minimize harm, justify aggression, or shift responsibility in ways that allow harmful behavior to continue without strong self-condemnation (Luo & Bussey, 2023).

This point is particularly relevant in relation to Indonesian school contexts. Recent evidence from Indonesia shows that negative social norms significantly predict bullying perpetration and victimization (Zaneva et al., 2025). This suggests that bullying prevention cannot rely only on individual moral reminders. If harmful speech is normalized within peer interaction, students may need collective norms and adult guidance to support ethical communication. In this sense, Socrates' Three Filters may help students evaluate speech, but they must be embedded within a wider school culture that consistently challenges humiliation, rumor, exclusion, and verbal aggression.

Empathy is another important support for the Three Filters. Without empathy, students may interpret the filter of goodness superficially, for example by assuming that a comment is acceptable as long as it is presented as a joke or as honesty. Empathy allows students to consider how words affect the feelings, dignity, and social position of others. Recent research identifies empathy as a protective factor in bullying contexts because it supports prosocial behavior and reduces tolerance toward harmful peer interaction (Hikmat et al., 2024). Similarly, Johander et al. (2022) found that teacher messages combining condemnation of bullying with empathy-raising elements can strengthen students' intention to stop bullying. These findings suggest that the Three Filters should be taught not only as cognitive questions, but also as prompts for perspective-taking.

Social-emotional learning also provides an important educational foundation. The three filters are closely related to several SEL competencies. The truth filter supports responsible decision-making and critical judgment. The goodness filter aligns with social awareness, empathy, and relationship skills. The usefulness filter supports self-management, problem-solving, and responsible communication. Fredrick et al. (2022) show that social-emotional learning practices are meaningfully related to bullying prevention because they strengthen students' interpersonal and self-regulation capacities. Therefore, the Three Filters become more actionable when they are connected to explicit SEL practices rather than taught as isolated moral advice.

School climate further determines whether reflective communication can become a shared habit. A positive school climate promotes belonging, safety, prosocial relationships, and consistent behavioral expectations, all of which are important for reducing bullying and aggression (Kelly et al., 2024; La Salle-Finley et al., 2024). Indonesian evidence also suggests that supportive family and school climate can protect children's subjective well-being in the context of bullying victimization (Borualogo et al., 2025). This means that reflective speech must be supported by relational and institutional consistency. A student may learn to ask whether a statement is true, good, and useful, but such reflection is more likely to influence behavior when teachers, peers, and school norms reinforce the same expectations.

Whole-school and restorative approaches also strengthen the practical relevance of the Three Filters. Reviews of school-based anti-bullying programs show that whole-school interventions can reduce bullying perpetration and victimization, although

effectiveness depends on implementation quality and school context (Gaffney et al., 2021a, 2021b; Pearce et al., 2024). Restorative approaches are also relevant because they emphasize accountability, dialogue, relational repair, and peaceful conflict resolution after harm occurs (Lodi et al., 2021; Pérez-Jorge et al., 2023). These approaches complement the Three Filters by helping students understand not only what they should avoid saying, but also how to repair harm when speech has already injured another person.

d. Proposed Conceptual Framework

Based on the integration of survey findings and literature review, this study proposes an integrated framework for using Socrates’ Three Filters in school-based bullying prevention. The framework positions truth, goodness, and usefulness as three interconnected dimensions of ethical speech. Truth addresses the problem of rumor, exaggeration, and unverified social labeling. Goodness addresses the problem of humiliation, disrespect, and the misuse of honesty as a justification for harm. Usefulness addresses the problem of purposeless or performative speech that produces embarrassment, exclusion, or peer approval at another student’s expense.

However, the framework also recognizes that individual reflection is not enough. The Three Filters become educationally meaningful only when supported by school-based practices. Empathy-building gives emotional depth to the filter of goodness. Social-emotional learning strengthens self-management, social awareness, relationship skills, and responsible decision-making. Teacher-guided intervention provides students with a consistent language for responding to harmful speech. Restorative dialogue helps students understand the impact of their words and repair relational harm. Whole-school expectations ensure that respectful communication becomes part of the school climate rather than an isolated classroom slogan.

Table 6. Conceptual Framework for Integrating Socrates’ Three Filters into School-Based Bullying Prevention

Component	Function in Ethical Speech	School-Based Support	Expected Educational Contribution
Truth	Encourages students to verify whether a statement is factual before repeating it	Critical questioning, teacher-guided dialogue, classroom discussion of rumor and misinformation	Reduces the spread of rumors, exaggeration, and harmful social labeling
Goodness	Encourages students to consider whether speech respects dignity and empathy	Empathy-building, perspective-taking, social-emotional learning	Reduces humiliating, degrading, and hostile communication
Usefulness	Encourages students to ask whether speech has a constructive purpose	Restorative dialogue, conflict reflection, responsible decision-making practice	Reduces impulsive comments and strengthens problem-solving communication
Integrated school practice	Embeds the filters into routines, norms, and adult responses	Whole-school expectations, consistent teacher intervention, peer leadership	Builds a respectful school climate and supports bystander awareness

This conceptual framework clarifies the contribution of the study. The Three Filters are not proposed as a complete anti-bullying theory or as a replacement for existing school-based interventions. Rather, they are proposed as a micro-ethical framework that can strengthen the communicative dimension of bullying prevention. Their value lies in translating abstract moral expectations into simple questions that students and teachers can use in everyday interaction: Is this true? Is this good? Is this useful?

e. School-Based Implications

The proposed framework has several practical implications. First, teachers can use the Three Filters as reflective prompts during classroom discussion, conflict mediation, advisory sessions, or restorative conversations. For example, when harmful language occurs, teachers can ask students whether the statement was true, whether it respected the dignity of others, and whether it served a constructive purpose. This shifts teacher response from merely punitive correction to moral and communicative guidance.

Second, the Three Filters can be integrated into social-emotional learning activities. The truth filter can be linked to critical thinking and responsible communication, the goodness filter to empathy and social awareness, and the usefulness filter to responsible decision-making and relationship repair, all of which are consistent with social-emotional learning and moral education perspectives (Fredrick et al., 2022; Mammen & Paulus, 2023). In this way, the filters become part of a broader educational process that supports both moral reasoning and interpersonal skill development.

Third, schools can incorporate the Three Filters into classroom agreements, peer leadership programs, anti-bullying campaigns, and whole-school expectations for respectful speech. This is important because bullying prevention requires consistency across classrooms, teachers, peer groups, and school routines. When the same language of truth, goodness, and usefulness is used across the school, students may be more likely to recognize harmful speech and understand how to respond to it.

Finally, the framework should be applied with caution. It should not be used to place responsibility on victims to tolerate harmful speech or to solve bullying by themselves. Bullying prevention remains the responsibility of adults, teachers, school leaders, and the wider school community. The Three Filters should be used to educate all students, especially potential bystanders and peer groups, while schools continue to provide protection, intervention, counseling, and consistent anti-bullying policies.

Overall, the contribution of this study is conceptual and pedagogical rather than causal or statistically generalizable. The student data provide exploratory evidence that harmful speech may involve uncertainty about truth, ambiguity between honesty and kindness, and limited reflection on usefulness. The literature explains why these patterns require broader support from empathy-building, social-emotional learning, teacher intervention, restorative dialogue, and positive school climate. Therefore, Socrates' Three Filters are best understood as a micro-ethical framework for reflective communication that can complement, but not replace, comprehensive school-based bullying prevention.

4. Conclusions

This exploratory study shows that Socrates' Three Filters of truth, goodness, and usefulness can serve as a practical micro-ethical framework for strengthening reflective communication in school-based bullying prevention. The descriptive survey findings indicate that 9 out of 10 students had witnessed or experienced bullying, 8 out of 10 were unsure whether information shared in bullying-related comments was true, 5 out of 10 were unsure whether they usually considered the necessity of their words before speaking, and 6 out of 10 believed that applying the Three Filters could help reduce bullying at school. These findings suggest that harmful student communication is not only related to intentional aggression, but also to weak verification habits, ambiguity between honesty and kindness, limited reflection on the purpose of speech, and uncertainty in responding to harmful peer interaction. The main contribution of this study lies in positioning Socrates' Three Filters as a conceptual-pedagogical framework for ethical speech rather than as a stand-alone anti-bullying solution. The novelty of the study is its integration of a classical philosophical framework with contemporary school-based bullying prevention, particularly by connecting truth, goodness, and usefulness with empathy-building, social-emotional learning, teacher-guided intervention, restorative dialogue, and whole-school expectations for respectful communication. Practically, the Three Filters can help teachers and schools guide students to ask whether their words are true, good or respectful, and useful or constructive before speaking or responding to harmful communication. These questions may be used in classroom discussions, conflict mediation, restorative conversations, social-emotional learning activities, and school-wide campaigns on respectful communication. However, the framework should not be used to place responsibility on victims or to replace formal anti-bullying policies. This study is limited by its small sample of 10 students from one school context, its reliance on self-reported perceptions, and its non-intervention design. Therefore, the findings should be interpreted as exploratory and conceptual rather than causal or statistically general-

izable. Future research should examine the application of Socrates' Three Filters through larger mixed-methods studies, classroom-based interventions, or longitudinal designs that measure changes in verbal aggression, bystander behavior, empathy, moral reasoning, and school climate.

References

- Bandura, A. (1977). *Social Learning Theory*. Prentice Hall.
- Borualogo, I. S., González-Carrasco, M., & Casas, F. (2024). Examining Predictors of Bullying Victimization in Indonesian Children. *Applied Research in Quality of Life*, 19(6), 3377–3405. <https://doi.org/10.1007/s11482-024-10383-0>
- Borualogo, I. S., Varela, J. J., & de Tezanos-Pinto, P. (2025). Sibling and School Bullying Victimization and Its Relation With Children's Subjective Well-Being in Indonesia: The Protective Role of Family and School Climate. *Journal of Interpersonal Violence*, 40(5–6), 1433–1458. <https://doi.org/10.1177/08862605241259412>
- Brickhouse, T. C., & Smith, N. D. (1990). *Socrates on trial* (2. print., with corr). Princeton Univ. Pr.
- Chiracu, A., Guțu, S., Golu, F., & Buică-Belciu, C. (2024). Bullying and Performance. The Role of School Climate. *Review of Psychopedagogy*, 13(1), 29–41. <https://doi.org/10.56663/rop.v13i1.72>
- Fredrick, S. S., Traudt, S., & Nickerson, A. B. (2022). Social Emotional Learning Practices in Schools and Bullying Prevention. In *Social Emotional Learning Practices in Schools and Bullying Prevention*. Routledge. <https://doi.org/10.4324/9781138609877-REE171-1>
- Gaffney, H., Ttofi, M. M., & Farrington, D. P. (2021a). Effectiveness of school-based programs to reduce bullying perpetration and victimization: An updated systematic review and meta-analysis. *Campbell Systematic Reviews*, 17(2), cl2.1143. <https://doi.org/10.1002/cl2.1143>
- Gaffney, H., Ttofi, M. M., & Farrington, D. P. (2021b). What works in anti-bullying programs? Analysis of effective intervention components. *Journal of School Psychology*, 85, 37–56. <https://doi.org/10.1016/j.jsp.2020.12.002>
- Hikmat, R., Suryani, S., Yosep, I., & Jeharsae, R. (2024). Empathy's Crucial Role: Unraveling Impact on Students Bullying Behavior - A Scoping Review. *Journal of Multidisciplinary Healthcare, Volume 17*, 3483–3495. <https://doi.org/10.2147/JMDH.S469921>
- Johander, E., Trach, J., Turunen, T., Garandeau, C. F., & Salmivalli, C. (2022). Intention to Stop Bullying following a Condemning, Empathy-Raising, or Combined Message from a Teacher – Do Students' Empathy and Callous-Unemotional Traits Matter? *Journal of Youth and Adolescence*, 51(8), 1568–1580. <https://doi.org/10.1007/s10964-022-01613-5>
- Kapile, C., Nuraedah, N., Junarti, J., & Nugroho, F. (2023). Bullying and its Implications on Middle School Students and Teachers in Indonesia. *AL-ISHLAH: Jurnal Pendidikan*, 15(1), 813–822. <https://doi.org/10.35445/alishlah.v15i1.2205>
- Kapitanoff, S., & Pandey, C. (2024). The Content of Verbal Bullying and Emotional Reactions Among Middle-School Students. *Child & Youth Care Forum*, 53(6), 1311–1332. <https://doi.org/10.1007/s10566-024-09796-7>
- Kelly, K. M., Riffle, L. N., Rodriguez-Harris, D., Wiemer, K., Jeong, R., Ayala, K. M., Demaray, M. K., & Malecki, C. K. (2024). The Moderation of School Climate on the Associations of Bullying Participant Behavior and GPA, Attendance, Office Disciplinary Referrals. *International Journal of Bullying Prevention*, 6(4), 373–388. <https://doi.org/10.1007/s42380-022-00148-4>
- La Salle-Finley, T., Yang, C., Espelage, D., & Jimerson, S. R. (2024). Understanding and Promoting School Climate, Bullying, and Social-Emotional Learning: Transdisciplinary and Transnational Science Advancing Positive Youth Outcomes. *School Psychology Review*, 53(5), 417–424. <https://doi.org/10.1080/2372966X.2024.2386235>
- Lodi, E., Perrella, L., Lepri, G. L., Scarpa, M. L., & Patrizi, P. (2021). Use of Restorative Justice and Restorative Practices at School: A Systematic Literature Review. *International Journal of Environmental Research and Public Health*, 19(1). <https://doi.org/10.3390/ijerph19010096>
- Luo, A., & Bussey, K. (2023). Moral disengagement in youth: A meta-analytic review. *Developmental Review*, 70, 101101. <https://doi.org/10.1016/j.dr.2023.101101>
- Mammen, M., & Paulus, M. (2023). The communicative nature of moral development: A theoretical framework on the emergence of moral reasoning in social interactions. *Cognitive Development*, 66, 101336. <https://doi.org/10.1016/j.cogdev.2023.101336>

- Meyer, K. (2024). Moral Education Through the Fostering of Reasoning Skills. *Ethical Theory and Moral Practice*, 27(1), 41–55. <https://doi.org/10.1007/s10677-023-10367-3>
- Olweus, D. (2013). *Bullying at School: What We Know and What We Can Do*. John Wiley & Sons.
- Pearce, N., Monks, H., Alderman, N., Hearn, L., Burns, S., Runions, K., Francis, J., & Cross, D. (2024). 'It's All About Context': Building School Capacity to Implement a Whole-School Approach to Bullying. *International Journal of Bullying Prevention*, 6(1), 53–68. <https://doi.org/10.1007/s42380-022-00138-6>
- Pérez-Jorge, D., Alonso-Rodríguez, I., Arriagada-Venegas, M., & Ariño-Mateo, E. (2023). Restorative methods as a strategy for the prevention of violence and bullying in primary and secondary schools in Mexico: An exploratory study. *Heliyon*, 9(7). <https://doi.org/10.1016/j.heliyon.2023.e18267>
- Ramadhanti, R., & Hidayat, M. T. (2022). Strategi Guru dalam Mengatasi Perilaku Bullying Siswa di Sekolah Dasar. *Jurnal Basicedu*, 6(3), 4566–4573. <https://doi.org/10.31004/basicedu.v6i3.2892>
- Tas'adi, R., & Fanany, R. (2023). *Defining Bullying in the Indonesian Context: A Systematic Review*. 421–431. https://doi.org/10.2991/978-2-38476-142-5_40
- UNESCO. (2025). *Behind the numbers: Ending school violence and bullying*. UNESCO. <https://doi.org/10.54675/TRVR4270>
- Vlastos, G. (1991). *Socrates, Ironist and Moral Philosopher*. Cornell University Press.
- Zaneva, M., Minnick, E., Nahar, Ginting, V., Aryani, F., Ohan, F., Haryanti, R. H., Winarna, S., Arsianto, Y., Budiawati, H., Widowati, E., Saraswati, R., Kristianto, Y., Suryani, Y. E., Ulum, D. F., & Bowes, L. (2025). Social Norms Predict Bullying: Evidence from an Anti-Bullying Intervention Trial in Indonesia. *International Journal of Bullying Prevention*, 7(4), 253–265. <https://doi.org/10.1007/s42380-023-00176-8>