

Moderation System

Complete guide to comment moderation, content reporting, user moderation, spam filtering, and audit logging

Version 1.1 | Updated 2026-04-11 | Moderators, Admin Users, Operations Teams, QA, Customer Success

[1. Simple Feature Overview](#)

The **Moderation System** lets your team review, approve, reject, and manage reader comments and interactions across the Hyphen platform. It protects your publication from spam, abusive content, and inappropriate behavior — while keeping the commenting experience smooth for genuine readers.

[What You Can Do](#)

Capability	What It Means
Comment Queue	See all reader comments waiting for review in one place
Approve / Reject / Flag / Hide	Take action on individual comments with a single click
Bulk Actions	Approve, reject, or mark as spam dozens of comments at once
Content Reports	Review reports submitted by readers about inappropriate comments
User Moderation	Warn, ban, or unban readers who repeatedly break the rules
Profanity Filter	Automatically catch comments containing blocked words
Spam Detection	Automatically score and flag comments that look like spam
Moderation Settings	Configure approval rules, comment length, link/image policies, and filter thresholds
Audit Trail	Every moderation action is logged with who did it, when, and why

[How Moderation Works — The Big Picture](#)

Reader writes a comment on an article (Reader Portal)

↓

System checks the comment against profanity and spam filters

↓

If "Require Approval" is ON → Comment goes to Pending status (not visible to other readers)

If "Require Approval" is OFF and filters pass → Comment is auto-Approved (visible immediately)

If filters flag the comment → Comment goes to Pending status for moderator review

↓

Moderator reviews the comment in the Admin Console Moderation page

↓

Moderator Approves → Comment becomes visible on Reader Portal

Moderator Rejects / Hides / Marks as Spam → Comment stays hidden from readers

↓

All actions are logged in the moderation log and audit trail

2. Who Should Use This Feature

Role	What They Do
Moderators	Review the comment queue daily, approve or reject comments, handle reports, warn users
Admin Users	Full access including moderation settings, profanity list management, and banning users
Operations Teams	Monitor moderation stats, ensure the queue is being cleared, check audit trails
QA Teams	Verify that moderation rules are working correctly, test comment flows end to end
Customer Success	Look up specific reader moderation history when handling complaints

[3. Before You Begin](#)

Before you can use the moderation features, make sure the following are in place:

[Required Permissions](#)

You need specific permissions assigned to your admin account. Ask an Admin to check your role in **Settings > Admin Users**.

Permission	What It Allows
moderation:read	View the Moderation page, comment queue, reports, user moderation records, and stats
moderation:approve	Approve comments, flag comments, restore comments, edit comment content
moderation:reject	Reject comments, mark as spam, hide comments, delete comments
moderation:ban_user	Warn users, ban users, unban users
settings:update	Access the Settings tab to change moderation configuration and manage the profanity word list

Typical role assignments:

- **Admin** role: All permissions above
- **Moderator** role: moderation:read, moderation:approve, moderation:reject
- **Viewer / Analyst** role: moderation:read only (view but cannot take action)

[Required Platform Configuration](#)

Prerequisite	Where to Check	What to Verify
Comments feature flag is enabled	Platform feature flags	<code>featureFlags.comments</code> must be <code>true</code>
Moderation settings are configured	Moderation > Settings tab	At minimum, confirm "Comments Enabled" is on
At least one article exists	Content section	Comments can only be posted on published articles
Reader accounts exist	Readers section	Readers must be registered and logged in to comment

4. Key Terms in Simple Language

Term	What It Means
Pending	A comment that is waiting for a moderator to review it. Readers cannot see pending comments.
Approved	A comment that has been reviewed and made visible to all readers on the article.
Rejected	A comment that was reviewed and turned down. It is hidden from readers permanently unless restored.
Flagged	A comment that has been marked for closer review — either by a moderator or by the automated filters. Still hidden from readers.
Spam	A comment identified as junk, advertisements, or automated spam. Hidden from readers.
Hidden	A comment that was manually hidden by a moderator. It exists in the system but readers cannot see it.
Report	When a reader clicks "Report" on a comment they think is inappropriate. Reports go to the Reports tab for moderator review.
Profanity Filter	An automated check that scans comment text against a list of blocked words you maintain.
Spam Filter	An automated check that scores comments for spam signals like excessive links, repeated text, or ALL CAPS.
Trusted User	A reader who has had enough comments approved (based on your threshold setting) to potentially bypass manual approval.
Moderation Log	A record of every action taken on a comment, report, or user — including who took the action and when.
Rate Limit	Readers are limited to 10 comments per 15 minutes and 5 reports per 15 minutes to prevent abuse.

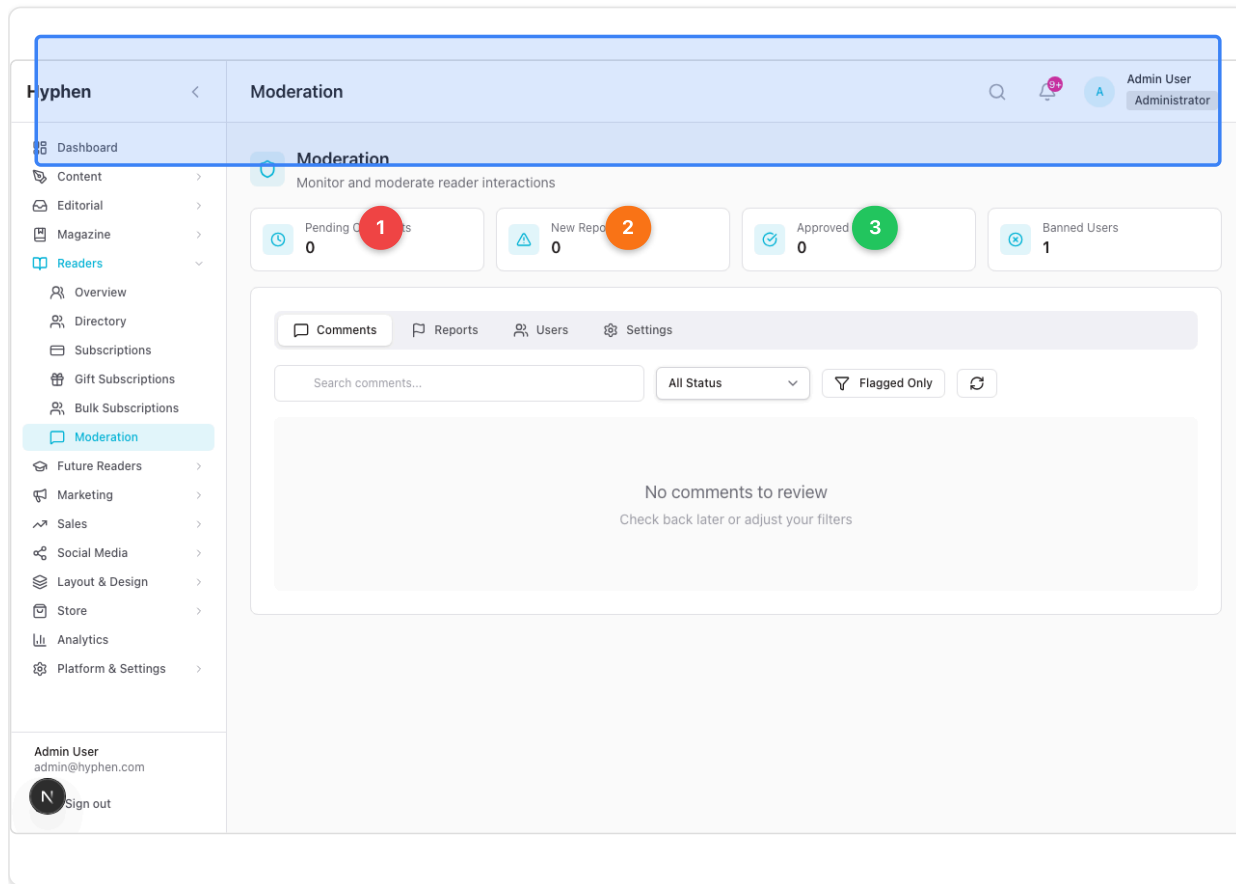
5. Step-by-Step Setup Guide

5.1 Accessing the Moderation Page

1. Log in to the **Admin Console**.
2. In the left sidebar, navigate to **Readers > Moderation**.
3. The Moderation page opens with a shield icon, titled "**Moderation**" with the subtitle "Monitor and moderate reader interactions."
4. If you see an "Access Denied" message, your account does not have the `moderation:read` permission. Contact an Admin to update your role.

"Note: There is also an older route at `/moderation` which automatically redirects to the current location at `/readers/moderation`."

5.2 Understanding the Moderation Dashboard



The moderation dashboard — your daily moderation overview

- Stat cards — pending comments, reported content, flagged users, and total actions at a glance
- 1 Comments tab — review and moderate reader comments
- 2 Reports tab — handle content and user reports
- 3 Users tab — manage warned or banned users

At the top of the Moderation page, you will see four **stat cards** giving you a quick snapshot:

Stat Card	Icon	What It Shows
Pending Comments	Clock	Number of comments waiting for review. Turns yellow/warning if more than 10 are waiting.
New Reports	Alert Triangle	Number of comments that have been reported by readers. Turns red if any reports exist.
Approved Today	Check Circle	Number of comments approved since midnight today.
Banned Users	X Circle	Number of currently banned readers.

Below the stat cards, you will see **four tabs**:

Tab	What It Contains
Comments	The comment moderation queue — your primary daily workspace
Reports	Reader-submitted reports about inappropriate comments
Users	Reader moderation records — warnings, bans, comment history
Settings	Moderation configuration — approval rules, filters, profanity words (requires settings:update permission)

5.3 Configuring Moderation Settings

“Permission required: `settings:update`”

Click the **Settings** tab on the Moderation page. You will see several groups of settings:

Comment Settings

Setting	What It Controls	Default
Comments Enabled	Master toggle — turns commenting on or off across the entire platform	On
Require Approval	When on, all new comments start as "Pending" and must be approved by a moderator before readers can see them. When off, comments go live immediately (unless flagged by filters).	On
Auto-approve for Trusted Users	When on, readers who have reached the trusted user threshold can have their comments auto-approved	On
Auto-close Comments After (days)	Automatically close comments on articles older than this many days. Set to 0 to never auto-close.	30
Minimum Comment Length	Shortest allowed comment in characters	10
Maximum Comment Length	Longest allowed comment in characters	2,000

Links & Media

Setting	What It Controls	Default
Allow Links	Whether readers can include URLs in comments	Off
Allow Images	Whether readers can include images in comments	Off

Profanity Filter

Setting	What It Controls	Default
Profanity Filter Enabled	Master toggle for the profanity filter	On
Profanity Action	What happens when a blocked word is detected: Flag (send to queue for review), Auto-reject (reject automatically), or Censor (replace the word with asterisks and allow the comment)	Flag

[Spam Filter](#)

Setting	What It Controls	Default
Spam Filter Enabled	Master toggle for the spam filter	On
Spam Threshold	Sensitivity level from 0 (catch everything) to 100 (catch almost nothing). Lower numbers mean stricter filtering.	70
Link Spam Threshold	Maximum number of links allowed in a single comment before it is flagged as spam	2

[User Trust Settings](#)

Setting	What It Controls	Default
Trusted User Threshold	Number of approved comments a reader needs before being considered "trusted"	10

[Notifications](#)

Setting	What It Controls	Default
Notify on Flag	Send a notification when a comment is automatically flagged by filters	On
Notify on Report	Send a notification when a reader reports a comment	On

To change any setting:

1. Go to the **Settings** tab.
2. Find the setting you want to change.
3. Toggle the switch, enter a new value, or select from the dropdown.
4. Click **Save** (or the setting may save automatically depending on the field).
5. Changes take effect immediately for all new comments.

[5.4 Managing the Profanity Word List](#)

“Permission required: `settings:update`”

The profanity word list is at the bottom of the **Settings** tab. This is a database-backed list of words that the profanity filter checks against.

[Adding a Word](#)

1. In the **Settings** tab, scroll to the **Profanity Words** section.
2. Type the word into the text field.
3. Select a **severity level** from the dropdown:
 - **Low** — Usually just flagged for review

- **Medium** — Can trigger flag or censor depending on your Profanity Action setting
- **High** — Typically auto-rejected or censored

4. Click **Add**.

5. The word appears in the list below. All words are stored in lowercase.

[Removing a Word](#)

1. Find the word in the profanity list (you can search/scroll).
2. Click the **Delete** or **Remove** button next to it.
3. The word is removed immediately.

“Tip: Start with high-severity words that should always be caught, then add medium and low severity words as you learn what your community encounters.”

[5.5 Reviewing Comments in the Queue](#)

Click the **Comments** tab. This is where you spend most of your moderation time.

[Filtering the Queue](#)

At the top of the Comments tab, you have three filter controls:

Filter	How to Use It
Search	Type text to search by comment content, author name, or article title
Status dropdown	Filter by: All, Pending, Approved, Rejected, Flagged, Spam. Start with Pending to see what needs review.
Flagged Only toggle	When turned on, shows only comments that have been flagged by filters or moderators

[What Each Comment Card Shows](#)

Each comment in the queue is displayed as a card with:

- **Article title and slug** — Which article the comment was posted on
- **Author name** — The reader's display name
- **Date** — When the comment was submitted
- **Status badge** — Color-coded: Pending (yellow), Approved (green), Rejected (red), Flagged (orange), Spam (gray), Hidden (gray)
- **Report count** — How many times readers have reported this comment
- **Comment text** — The full content of the comment
- **Action buttons** — The actions you can take (see next section)

Pagination

The queue shows 20 comments per page. Use the **Previous** and **Next** buttons at the bottom to navigate. You'll see "Page X of Y" to track your position.

Click the **Refresh** button to reload the queue with the latest comments.

5.6 Taking Action on Individual Comments

The screenshot displays the Hyphen Moderation interface. On the left is a navigation sidebar with the following items: Dashboard, Content, Editorial, Magazine, Readers, Overview, Directory, Subscriptions, Gift Subscriptions, Bulk Subscriptions, Moderation (highlighted), Future Readers, Marketing, Sales, Social Media, Layout & Design, Store, Analytics, and Platform & Settings. The main content area is titled 'Moderation' and includes a search bar, a user profile for 'Admin User Administrator', and four summary cards: Pending Comments (0), New Reports (0), Approved Today (0), and Banned Users (1). Below these are tabs for 'Comments', 'Reports', 'Users', and 'Settings'. A search bar for comments is present, along with a status filter set to 'All Status' and a 'Flagged Only' button. The main area shows a message: 'No comments to review. Check back later or adjust your filters.'

Review pending comments with approve, reject, flag, and ban actions for each entry.

Each comment card has action buttons. The buttons available depend on the comment's current status and your permissions.

Action	What It Does	Required Permission	When to Use It
Approve	Sets status to Approved. Comment becomes visible on the Reader Portal.	moderation:approve	The comment is appropriate and should be published.
Reject	Sets status to Rejected. Comment is hidden from readers.	moderation:reject	The comment violates your guidelines and should not be published.
Flag	Sets status to Flagged. Comment stays hidden, marked for further review.	moderation:approve	You want to escalate the comment for another moderator or admin to review.
Spam	Sets status to Spam. Comment is hidden and categorized as junk.	moderation:reject	The comment is clearly automated spam or junk.
Hide	Sets status to Hidden. Comment is hidden from readers but kept in the system.	moderation:reject	You want to temporarily remove a comment without fully rejecting it.
Restore	Sets status back to Pending. Comment re-enters the review queue.	moderation:approve	You or another moderator made an error and the comment needs re-review.
Edit	Opens the comment content for editing.	moderation:approve	The comment is mostly fine but contains a minor issue (e.g., a phone number or personal info that should be removed).
Delete	Permanently removes the comment from the system.	moderation:reject	The comment should be completely removed. This cannot be undone.
View User	Opens the User Moderation Panel for the comment author.	moderation:read	You want to see this reader's moderation history, other comments, or take action on their account.

To take an action:

1. Find the comment in the queue.
2. Review the comment text, author, and article context.
3. Click the appropriate action button.
4. For some actions, you may be prompted to add a **moderation note** explaining your decision.
5. The comment's status updates immediately.

“Important: When you approve a comment, the article's comment count increases by one. When you reject, hide, spam, or delete a previously approved comment, the count decreases by one.”

[5.7 Using Bulk Actions on Multiple Comments](#)

When you have many comments to process at once (e.g., after a weekend), bulk actions save time.

1. In the **Comments** tab, you will see a **checkbox** next to each comment.
2. Use the **Select All** checkbox at the top to select all comments on the current page, or check individual comments.

3. Once at least one comment is selected, three bulk action buttons appear:
 - **Approve** — Approve all selected comments
 - **Reject** — Reject all selected comments
 - **Spam** — Mark all selected comments as spam
4. Click the desired action.
5. The system processes all selected comments. You will see a result showing how many succeeded and how many failed (if any).

“Limit: You can bulk-moderate up to 100 comments at once.”

5.8 Managing Content Reports

Click the **Reports** tab to see reports submitted by readers.

[What a Report Shows](#)

Each report includes:

Field	What It Means
Report ID	Unique identifier
Type	What kind of report (content or user)
Status	Open, Reviewing, Resolved, or Dismissed
Reason	Why the reader reported it: Spam, Harassment, Hate Speech, Misinformation, Copyright, Inappropriate, or Other
Target	The comment content (or article/user) that was reported
Reporter	The reader who submitted the report
Date	When the report was submitted

[Report Statuses Explained](#)

Status	What It Means
Open	New report that has not been looked at yet
Reviewing	A moderator has started looking into this report
Resolved	The report has been addressed — the moderator took action (e.g., rejected the comment)
Dismissed	The report was reviewed and determined to be invalid or no action was needed

[Taking Action on Reports](#)

For each report, you can:

1. **Update Status** — Change from Open to Reviewing to indicate you are working on it.

2. **Resolve** — Mark the report as resolved. You will be prompted to add a resolution note explaining what action was taken.
3. **Dismiss** — Mark the report as dismissed. You will be prompted to add a reason why no action was needed.

Typical report workflow:

1. Open the **Reports** tab.
2. Filter by **Open** status to see new reports.
3. Click on a report to see the details.
4. Read the reported content and the reporter's reason.
5. If action is needed, go to the **Comments** tab and reject/hide/spam the reported comment.
6. Come back to the report and click **Resolve**, adding a note like "Comment rejected — violated community guidelines."
7. If no action is needed, click **Dismiss** and add a reason like "Comment does not violate guidelines."

"Note: Readers cannot see who reported a comment. Reporter identity is only visible to admin users."

5.9 Moderating Users (Warn, Ban, Unban)

Manage flagged users with warn, ban, and unban actions plus comment history review.

Click the **Users** tab to see a list of readers who have comment activity.

[What the Users Tab Shows](#)

The Users tab displays a table with:

Column	What It Shows
User	Reader's name
Status	Active, Warned, Restricted, or Banned
Comments	Total number of comments submitted
Warnings	Number of warnings issued to this reader
Reports	Number of reports filed against this reader
Actions	A View button to open the full User Moderation Panel

[Using the User Moderation Panel](#)

Click **View** next to any user to open the **User Moderation Panel** (a modal/slide-out). This shows:

- **User information** — Name, email, account details
- **Statistics** — Total comments, approved, rejected, flagged counts
- **Active warnings** — Any warnings currently on the account
- **Active bans** — Any bans currently in effect
- **Recent comments** — The user's latest comments

[Actions Available in the User Moderation Panel](#)

Action	What It Does	Required Permission
Warn	Issues a formal warning to the reader. The warning is logged.	moderation:ban_user
Ban	Bans the reader from commenting. You select the ban type: Comment ban, Platform ban, Temporary ban, or Permanent ban. You provide a reason.	moderation:ban_user
Unban	Removes an active ban, allowing the reader to comment again.	moderation:ban_user
Toggle Trusted	Marks or unmarks the reader as a trusted user.	moderation:ban_user

What happens when a reader is banned:

- The reader can no longer submit comments on any article.
- If they try to comment, they see a message: "Your account has been restricted from commenting."
- Their existing approved comments remain visible unless you separately reject or hide them.
- The ban reason is recorded in the system.

To ban a reader:

1. Open the **Users** tab.
2. Click **View** next to the reader.
3. In the User Moderation Panel, click **Ban**.
4. Select the ban type (e.g., Comment, Permanent).
5. Enter a reason for the ban.
6. Confirm the action.

To unban a reader:

1. Open the **Users** tab.
2. Find the banned reader (their status shows "Banned").
3. Click **View**.
4. Click **Unban**.
5. The reader can immediately start commenting again.

5.10 How Readers Submit Comments on the Reader Portal

Understanding the reader experience helps you moderate effectively.

1. A reader visits an article on the **Reader Portal**.
2. Below the article, they see the **comment section** (if commenting is enabled for that article).
3. The reader must be **logged in** to comment. If not logged in, they see a "Sign in" link.
4. They type their comment in the text box. The box shows:
 - Placeholder text: "Share your thoughts..."
 - A character counter (max 2,000 characters)
 - A **Comment** button
5. When they click **Comment**, the system:
 - Validates the comment length (minimum 10, maximum 2,000 characters)
 - Checks for banned words (profanity filter)
 - Checks for spam signals (spam filter)
 - Checks if the reader is banned
 - Checks the rate limit (max 10 comments per 15 minutes)
6. If the comment passes all checks:

- If **"Require Approval" is ON**: The comment is saved as Pending. The reader sees: "Your comment has been submitted and is awaiting moderation."
- If **"Require Approval" is OFF and filters pass**: The comment is saved as Approved and immediately visible to other readers.
- If **filters flag the comment**: The comment is saved as Pending regardless of the approval setting.

Replying to comments:

- Readers can click **Reply** on any approved comment to write a threaded reply.
 - Replies follow the same validation and moderation rules as top-level comments.
 - The reply box shows the placeholder "Write a reply..."
-

[5.11 How Readers Report Comments on the Reader Portal](#)

1. Next to each comment, readers see a **Report** button.
2. Clicking it opens the **Report Comment** modal with the message: "Help us keep the community safe..."
3. The reader selects a reason by clicking one of seven radio buttons:
 - Spam
 - Harassment
 - Hate Speech
 - Misinformation
 - Copyright
 - Inappropriate
 - Other
4. Optionally, the reader adds details in the "Additional details (optional)" text box (max 1,000 characters).
5. They click **Submit Report**.
6. The report is created with status "Open" and appears in your Reports tab.
7. The reported comment's report count increases by one.

Protections against report abuse:

- A reader can only have one active (Open or Reviewing) report against the same comment. Duplicate reports are blocked with an error.
 - Readers are limited to 5 reports per 15 minutes per IP address.
-

5.12 How Moderation Actions Affect Reader Portal Visibility

This is important to understand: **only comments with "Approved" status are visible on the Reader Portal.**

Comment Status	Visible on Reader Portal?	What the Reader Sees
Pending	No	The comment author sees a message that their comment is awaiting moderation. Other readers see nothing.
Approved	Yes	The comment appears under the article with the author's name, date, and content.
Rejected	No	The comment disappears from the article for all readers.
Flagged	No	Same as Pending — hidden from all readers while under review.
Spam	No	Completely hidden from all readers.
Hidden	No	Completely hidden from all readers.

Article comment count: The number shown next to the article's comment count reflects only Approved comments. When you approve a comment, the count goes up by one. When you reject, hide, or spam an approved comment, the count goes down by one.

5.13 Audit Logging and Moderation History

Every moderation action is automatically recorded in two places:

[Moderation Log](#)

The moderation log captures:

- **What action** was taken (approve, reject, flag, spam, hide, restore, delete, edit, warn, ban, unban, resolve report, dismiss report)
- **Which item** was affected (the comment, report, or reader)
- **Who** performed the action (the admin user)
- **Previous status** and **new status** (e.g., pending → approved)
- **Reason or note** provided by the moderator
- **When** the action occurred

[Audit Trail](#)

All moderation actions are also logged in the platform's general audit trail system, which records:

- The action type (e.g., moderation_approve, moderation_reject, moderation_ban_reader)
- The actor (admin user)
- IP address and user agent

- Timestamp

This dual logging means you have both a moderation-specific history and a platform-wide audit trail for compliance and accountability.

5.14 GDPR and Compliance Considerations

When handling moderation, keep these compliance points in mind:

Consideration	What to Know
Reporter privacy	When a reader reports a comment, their identity (reporter name/email) is visible only in the Admin Console. Readers on the Reader Portal cannot see who reported a comment.
Reader data in comments	If a reader requests data export (GDPR), their comments and moderation history are part of their data.
Account deletion	If a reader requests account deletion, consider how this affects their comments. Comments may need to be anonymized or removed.
Moderation notes	Notes you add to moderation actions are stored in the system. Avoid including unnecessary personal information in moderation notes.
Ban reasons	Ban reasons are stored in the reader's record. Keep them factual and policy-based (e.g., "Repeated violation of community guidelines — 3 rejected comments for harassment").
Audit retention	Moderation logs and audit entries are retained as part of the platform's compliance records.

6. How to Verify It Worked

After setting up or using moderation features, use these checks to confirm everything is working:

Settings Verification

What to Check	How to Check
Settings saved correctly	Go to Moderation > Settings tab, confirm your values are shown
Comments are enabled	Post a test comment on the Reader Portal — the comment form should appear
Require Approval is working	Post a test comment — it should appear as "Pending" in the queue, not visible to other readers
Profanity filter is working	Post a test comment containing a blocked word — it should be flagged or rejected based on your Profanity Action setting

Comment Moderation Verification

What to Check	How to Check
Approve works	Approve a pending comment → Check the Reader Portal — the comment should now appear under the article
Reject works	Reject an approved comment → Check the Reader Portal — the comment should disappear
Hide works	Hide an approved comment → Check the Reader Portal — the comment should disappear
Restore works	Restore a rejected comment → It should appear as Pending in your queue again
Bulk actions work	Select multiple comments → Click bulk Approve → All should change to Approved
Comment count updates	After approving/rejecting → Check the article's comment count on the Reader Portal

Report Verification

What to Check	How to Check
Reports appear	Have a test reader report a comment → Check the Reports tab — the report should appear as "Open"
Report resolution works	Resolve a report → Its status should change to "Resolved"
Duplicate prevention works	Have the same reader report the same comment again → They should see an error

User Moderation Verification

What to Check	How to Check
Ban works	Ban a test reader → Have them try to comment → They should see "Your account has been restricted from commenting"
Unban works	Unban the reader → They should be able to comment again
Warning logged	Warn a reader → Check their User Moderation Panel — the warning should appear

[7. Worked Examples](#)

[7.1 Example 1: Reviewing and Approving a New Comment](#)

Scenario: A reader named Priya posted a comment on an article titled "The Future of Urban Fiction." Your settings have "Require Approval" turned on, so the comment is waiting in the queue.

Steps:

1. Log in to the **Admin Console**.
2. Navigate to **Readers > Moderation**.
3. Check the stat cards — you see **Pending Comments: 3**, confirming there are comments waiting.
4. Click the **Comments** tab (it should be selected by default).
5. Set the **Status** dropdown to **Pending** to focus on comments needing review.
6. Find Priya's comment card. You can see:
 - Article: "The Future of Urban Fiction"
 - Author: Priya
 - Status: **Pending** (yellow badge)
 - Content: "This article really resonated with me. The parallels between urban fiction and real city life are brilliantly explored."
7. The comment is thoughtful and appropriate. Click the **Approve** button.
8. The status badge changes to **Approved** (green).
9. **Verify on Reader Portal:** Open the article "The Future of Urban Fiction" on the Reader Portal. Scroll to the comments section. Priya's comment should now be visible with her name, the date, and the comment text. The article's comment count should have increased by one.

[7.2 Example 2: Rejecting Inappropriate Content and Verifying Reader Portal](#)

Scenario: A reader named Vikram posted a comment containing offensive language on a poetry article. The profanity filter flagged it, but you need to review and take final action.

Steps:

1. Navigate to **Readers > Moderation**.
2. Notice the **Pending Comments** stat card shows pending items.
3. In the **Comments** tab, toggle **Flagged Only** to **On** to see only flagged comments.
4. Find Vikram's comment. You can see:

- Status: **Pending** (it was flagged by the profanity filter)
 - Content contains offensive language
 - Report count may show 0 (it was caught by the filter, not reported by readers)
5. Click **Reject**.
 6. Optionally add a moderation note: "Contains offensive language — violates community guidelines."
 7. The status changes to **Rejected** (red badge).
 8. **Verify on Reader Portal:** Open the article on the Reader Portal. Vikram's comment should not appear anywhere in the comment section. The comment count should not include this comment.
 9. **Consider user action:** If Vikram has a pattern of inappropriate comments, click **View User** on his comment card to open the User Moderation Panel. Review his history. If warranted, click **Warn** or **Ban**.
-

[7.3 Example 3: Restoring a Comment That Was Moderated Incorrectly](#)

Scenario: A colleague accidentally rejected a perfectly fine comment from a reader named Anita. The customer success team received a complaint and asked you to fix it.

Steps:

1. Navigate to **Readers > Moderation**.
2. In the **Comments** tab, set the **Status** dropdown to **Rejected** to find Anita's comment.
3. Use the **Search** box to type "Anita" to narrow the results.
4. Find Anita's comment. Read the content to confirm it is indeed appropriate.
5. Click **Restore**.
6. The status changes to **Pending** (the comment re-enters the review queue).
7. Now click **Approve** to make the comment visible to readers.
8. The status changes to **Approved** (green).
9. **Verify on Reader Portal:** Open the article where Anita commented. Her comment should now appear in the comment section.
10. **Inform the team:** Let the customer success team know the comment has been restored and is now visible.

"Note: The moderation log will show the full history: original rejection, the restore action, and the approval — including who performed each action and when."

7.4 Example 4: Checking Moderation History for a Specific Comment

Scenario: Your editorial lead asks about a comment on a sensitive political article that was reportedly removed. They want to know who removed it, when, and why.

Steps:

1. Navigate to **Readers > Moderation**.
2. In the **Comments** tab, set the **Status** dropdown to **All** (to search across all statuses, including rejected and hidden).
3. Use the **Search** box to search for the article title or the commenter's name.
4. Find the comment in question. Note its current status (e.g., Rejected or Hidden).
5. The comment card shows:
 - Who moderated it (the "Moderated by" field)
 - When it was moderated (the moderation timestamp)
 - The moderation note (if the moderator added one)
6. For the full audit trail, the moderation log records every status change, including:
 - The previous status and new status
 - The admin user who took the action
 - The timestamp
 - Any reason provided
7. Report back to the editorial lead with the specific details: "The comment was rejected by [Moderator Name] on [Date] with the note: [reason]."

"Tip: If you need to verify this through the platform's broader audit system, check the Audit section which captures all moderation actions with IP address and timestamp details."

7.5 Example 5: Troubleshooting a Missing or Still-Visible Comment

Scenario A — Comment not showing in moderation queue:

A reader emails saying they posted a comment but it never appeared. You check the moderation queue and don't see it.

Troubleshooting steps:

1. **Check if commenting is enabled:** Go to **Moderation > Settings** tab. Confirm "Comments Enabled" is **On**.
2. **Check the platform feature flag:** Confirm `featureFlags.comments` is `true`.

3. **Check all statuses:** In the Comments tab, set Status to **All** and search for the reader's name. The comment might already have been auto-rejected by the spam or profanity filter.
4. **Check if the reader is banned:** Go to the **Users** tab, search for the reader. If their status is "Banned," their comment submission would have been blocked entirely (it never entered the system).
5. **Check rate limits:** If the reader posted many comments quickly (more than 10 in 15 minutes), subsequent comments would have been rejected with a rate limit error.
6. **Check comment length:** If the reader's comment was shorter than the minimum length (default 10 characters) or longer than the maximum (default 2,000 characters), it would have been rejected at submission time.

Scenario B — Comment still visible after rejection:

You rejected a comment but a reader reports it is still visible on the Reader Portal.

Troubleshooting steps:

1. **Verify the rejection:** Go to **Moderation > Comments**, search for the comment. Confirm its status is **Rejected** (not still Pending or Approved).
 2. **Check for caching:** The Reader Portal may be showing a cached version of the page. Ask the reader to refresh the page or try in an incognito/private browser window.
 3. **Check the correct article:** Confirm you rejected the comment on the correct article. The reader may be looking at a different article.
 4. **Check for duplicates:** The reader may have posted the same comment twice. Search for similar content in the queue — there may be a second copy that is still Approved.
-

8. Common Mistakes and How to Fix Them

Mistake	What Happens	How to Fix
Forgetting to set "Require Approval" to On	All comments go live immediately without moderator review. Inappropriate content appears on the Reader Portal.	Go to Moderation > Settings , turn on Require Approval . Existing approved comments are not affected, but all new comments will require approval.
Rejecting a comment instead of flagging it	The comment is permanently marked as rejected. If you meant to escalate it for senior review, it won't be obvious.	Click Restore on the rejected comment (this sets it back to Pending). Then click Flag to mark it for escalation.
Deleting a comment instead of rejecting it	The comment is permanently removed from the system and cannot be recovered.	Deletion is irreversible. In the future, use Reject or Hide instead of Delete unless you are certain the comment should be permanently removed.
Banning a user without checking their history	You may ban a user for one bad comment when they have a long history of good contributions.	Always click View User before banning. Review their total comments, approved count, and previous warnings. Issue a warning first for first-time offenders.
Not adding moderation notes	When another moderator or manager asks why a comment was rejected, there is no explanation on record.	Always add a brief moderation note when rejecting, hiding, or flagging a comment. Example: "Contains personal information — privacy concern."
Approving comments with personal information	A comment containing someone's phone number, email, or address becomes publicly visible.	Before approving, scan the comment for personal data. If found, use Edit to remove the personal information, then Approve .
Not processing reports	Open reports pile up. Readers lose trust that their reports are being reviewed.	Check the Reports tab regularly. Aim to review all Open reports within 24 hours. Set a daily reminder if needed.
Setting spam threshold too low	Legitimate comments get flagged as spam, creating extra work and frustrating readers.	Increase the Spam Threshold in Settings (higher number = less strict). Start at 70 and adjust based on false positive rate.
Profanity filter set to "Auto-reject" with a large word list	Comments get silently rejected and readers don't understand why.	Consider using Flag instead of Auto-reject as your Profanity Action, so a human reviews flagged comments. Only use Auto-reject for the most clearly offensive words.
Not unbanning a reader after the issue is resolved	The reader remains locked out of commenting even after the situation is addressed.	Set a reminder to review bans periodically. Go to the Users tab, find banned readers, and decide if any should be unbanned.

9. QA / Testing Checklist

Use this checklist to validate that the moderation system is working correctly.

Setup and Permissions

- Admin with `moderation:read` can access the Moderation page
- Admin without `moderation:read` sees "Access Denied" on the Moderation page
- Admin with `moderation:approve` can approve and restore comments
- Admin with `moderation:reject` can reject, flag, spam, hide, and delete comments
- Admin without `moderation:approve` cannot see or click the Approve button
- Admin with `moderation:ban_user` can warn and ban users
- Admin with `settings:update` can see and modify the Settings tab
- Admin without `settings:update` cannot see the Settings tab

Comment Submission (Reader Portal)

- Logged-in reader can submit a comment
- Comment form shows character counter (max 2,000)
- Comment shorter than minimum length (default 10) is rejected with an error
- Comment longer than maximum length (default 2,000) is rejected with an error
- Non-logged-in reader sees "Sign in" instead of the comment form
- Banned reader sees restriction message and cannot submit
- Submitting more than 10 comments in 15 minutes triggers rate limit error
- With "Require Approval" ON, new comments show as Pending in the admin queue
- With "Require Approval" OFF and clean content, comments appear immediately as Approved
- Reader sees "Your comment has been submitted and is awaiting moderation" for pending comments

Profanity and Spam Filters

- Comment containing a blocked word (high severity) is handled per Profanity Action setting (flagged, auto-rejected, or censored)
- Comment containing a blocked word (low severity) is flagged for review
- Comment with excessive links (above Link Spam Threshold) is flagged
- Comment with links is rejected if "Allow Links" is off
- Comment with images is rejected if "Allow Images" is off

- Adding a word to the profanity list works — new comments containing it are caught
- Removing a word from the profanity list works — new comments containing it pass through

Moderation Actions (Admin Console)

- Approving a pending comment changes status to Approved
- Approved comment appears on the Reader Portal under the correct article
- Rejecting a comment changes status to Rejected
- Rejected comment does not appear on Reader Portal
- Flagging a comment changes status to Flagged
- Marking as Spam changes status to Spam
- Hiding a comment changes status to Hidden
- Restoring a rejected comment changes status to Pending
- Editing a comment updates the content in the system
- Deleting a comment removes it permanently
- Article comment count increases when a comment is approved
- Article comment count decreases when an approved comment is rejected/hidden/deleted

Bulk Actions

- Selecting multiple comments and clicking bulk Approve changes all to Approved
- Selecting multiple comments and clicking bulk Reject changes all to Rejected
- Selecting multiple comments and clicking bulk Spam changes all to Spam
- Select All checkbox selects all comments on the current page

Content Reports

- Reader can report a comment from the Reader Portal
- Report modal shows all 7 reason options
- Report appears in the Reports tab with status "Open"
- Duplicate report by same reader on same comment is blocked
- More than 5 reports in 15 minutes triggers rate limit
- Moderator can update report status to "Reviewing"
- Moderator can resolve a report with a resolution note
- Moderator can dismiss a report with a reason

[User Moderation](#)

- Users tab shows readers with comment activity
- View button opens User Moderation Panel with correct statistics
- Warning a reader creates a logged warning
- Banning a reader prevents them from commenting
- Banned reader sees restriction message on Reader Portal
- Unbanning a reader restores their ability to comment

[Audit and Logging](#)

- Approving a comment creates a moderation log entry
- Rejecting a comment creates a moderation log entry
- Banning a user creates a moderation log entry
- All moderation actions appear in the platform audit trail
- Log entries include the actor, action, timestamp, and state change

[Dashboard Stats](#)

- Pending Comments count matches actual pending comments
 - New Reports count matches actual open reports
 - Approved Today count resets at midnight
 - Banned Users count matches actual banned readers
-

10. Current Known Limitations

The following limitations exist in the current implementation. They are noted here so you know what to expect and can plan workarounds where needed.

Limitation	What It Means	Workaround
Trusted user auto-approval not fully active	The setting "Auto-approve for Trusted Users" exists and can be toggled, but the system does not yet automatically approve comments from trusted users. All comments still follow the standard approval rules.	Continue manually approving comments from trusted users. The feature will be fully connected in a future update.
Auto-close comments not enforced	The "Auto-close Comments After X Days" setting can be configured, but commenting is not yet automatically disabled on older articles.	Manually disable comments on individual articles if needed, or monitor old articles for new comments.
Notification delivery not connected	The "Notify on Flag" and "Notify on Report" settings exist, but actual email or push notification delivery for these events is not yet wired up.	Check the moderation queue and reports tab regularly (at least daily) rather than relying on notifications.
Temporary ban duration not tracked	The ban type selector includes "Temporary" as an option, but there is no automatic expiration. A temporary ban stays active until manually removed.	If you issue a temporary ban, set a personal reminder to revisit and unban the reader after the intended period.
No comment edit history	When a moderator edits a comment, there is no visible "edited at" timestamp on the comment. The original and edited content are not displayed side by side.	Add a moderation note when editing a comment to record what was changed and why.
Comments are permanently deleted	The Delete action removes a comment from the database entirely. There is no soft-delete or trash/recycle bin.	Use Reject or Hide instead of Delete when you might need to reference the comment later. Only delete when absolutely necessary.
No per-comment moderation history view	There is no single view showing all moderation actions taken on a specific comment over time.	Check the moderation log filtered by the comment in question, or review the audit trail.
No bulk report management	You can bulk-moderate comments, but there is no bulk action for reports. Each report must be resolved or dismissed individually.	Process reports one at a time. Prioritize by focusing on "Open" status reports first.
No reader appeal process	Banned readers have no built-in way to submit an appeal or request unbanning through the Reader Portal.	Handle appeals through your existing customer support channel (email, support form). A moderator with <code>moderation:ban_user</code> permission can then unban the reader manually.

This guide is based on the Moderation module implementation as of March 2026. If you notice any behavior that differs from what is described here, please contact the platform team for clarification.

[< Magazine Issue Management](#)

[E-Commerce Store >](#)

[Complete guide to creating magazine issues, managing...](#)

[Complete guide to the Shopify-integrated e-commerce...](#)