

2 PROMPTS · 7 REFERENCES · 3 ROUTES

HOW TO MAKE A 2D STICK FIGURE HISTORY VIDEO WITH AI

The look of a 4.1M views history video, and how to check a video is worth copying first.



A history video from a channel with **8.59K subscribers** reached **4.1 million views in 8 days**. The whole look fits in three lines:

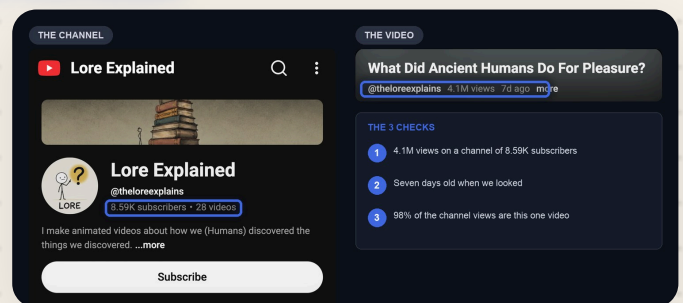
Stick figures with a thick black outline
Flat colored backgrounds, almost empty
A new picture every 2 seconds

We rebuilt its first minute: **34 scenes, 68 seconds**, generated in 9 minutes 42 seconds. This guide checks the video first, then gives you the prompts and the references to copy that look, and three ways to use them.

1 · CHECK THAT THE VIDEO IS A REAL OUTLIER

An outlier is a video with far more views than the rest of its channel. Check four things on YouTube, no tools needed:

- Views against subscribers.** Ours: 4.1M views, 8.59K subscribers
- Age.** 7 days old when we looked
- The rest of the channel.** The other 27 videos had a median of 1,183 views. This one held 98% of the channel's views
- The exact title, in quotes, in the search bar.** 43 videos had it: 19 before, 23 after. The best before got 917k. The 23 after got 122k together



REAL CHANNEL AND VIDEO PAGES · THE NUMBERS WE USED

Then check how it is made. We counted 30 pictures in 59 seconds, and the frames inside each one are identical. Nothing moves, so there is nothing to animate.

2

COPY THE PROMPTS FIRST

Everything is at the end of this guide: the **7 references** on page 4 and the **prompts** on page 5. Copy them, then pick a route.

WHAT YOU NEED

- The prompts and the 7 reference images, at the end of this guide
- A script for your video. Write your own, or use the narration of any video
- An image tool and a video editor such as CapCut, for routes A and B
- The video to compare with. Ours: youtube.com/watch?v=oQQRvdyFjM4

3 · ROUTES A AND B: CHATGPT AND CAPCUT, OR OUR PROMPT

Route A, the manual route. This is how most people do it. It works. It is also the slowest way.

- 1 Open ChatGPT. Attach the 7 references (page 4), or the one sheet on the web page. Paste the prompts too (page 5)
- 2 Send the message on the right, then paste your script
- 3 Generate every scene on its own with any image tool. Our first minute needed 34 images
- 4 Record your voiceover, or make it with any text to speech tool
- 5 Put the images and the voiceover in CapCut or any video editor. Cut each image to the length of its sentence
- 6 Export the video and upload it to YouTube

SEND THIS TO CHATGPT

These images and prompts show the look of a video. Write one image prompt for each scene of the script below, in that same look. Keep the main character identical in every scene.

Route B, the easy route. You do not need ChatGPT. The Visual Style on page 5 is already a fill in the blanks prompt. Put it at the start of every scene prompt and replace the words in [BRACKETS] with your scene. Then generate, join and upload as in route A. If your tool animates images, put the Video Style at the start of every motion prompt.

ONE SCENE, FILLED IN

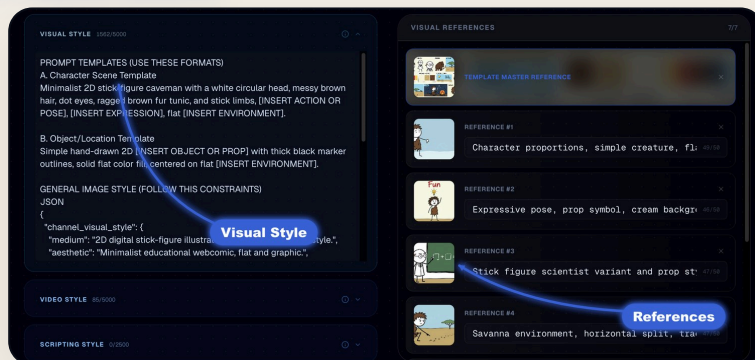
Minimalist 2D stick-figure caveman with a white circular head, messy brown hair, dot eyes, ragged brown fur tunic, and stick limbs, sitting on the ground with his chin on his hand, bored expression, flat cream environment.

4 · ROUTE C: PASTE THEM INTO A TEMPLATE

Making 34 images one by one is the slow part. In [Calliope](#) the prompts and the references live in one reusable template. Paste your script once and every scene comes out in the same look.

- 1 Open the **Visuals Lab**
- 2 Paste the **Visual Style** and the **Video Style** into their boxes. Both are editable
- 3 In **Visual References** add the 7 images and type each label (page 5) in its box. The panel should read 7/7
- 4 Click **Save Videos Template**

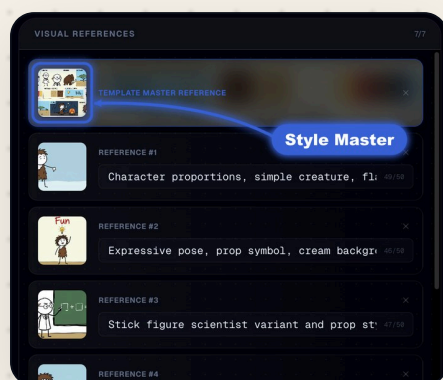
We did not write ours: Calliope read the look from the video's YouTube link in 72 seconds, with no edits. A style read from a link can come out wrong, so compare it with the video, as in step 5.



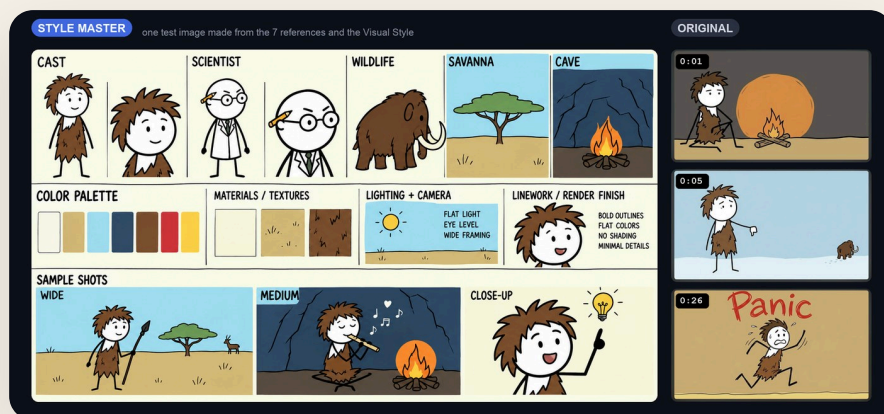
THE TEMPLATE · VISUAL STYLE LEFT, THE 7 REFERENCES RIGHT

5 • TEST THE STYLE BEFORE YOU GENERATE

Click **Test Style**. You get the **Style Master**: one sheet with the cast, the places, a palette and sample shots, made from your references and your Visual Style. It becomes the first reference of the template. Put it next to frames of the original. Ours read as the same family, so we went ahead. If yours looks like another channel, go back to step 4 and fix the texts.



THE STYLE MASTER, FIRST REFERENCE

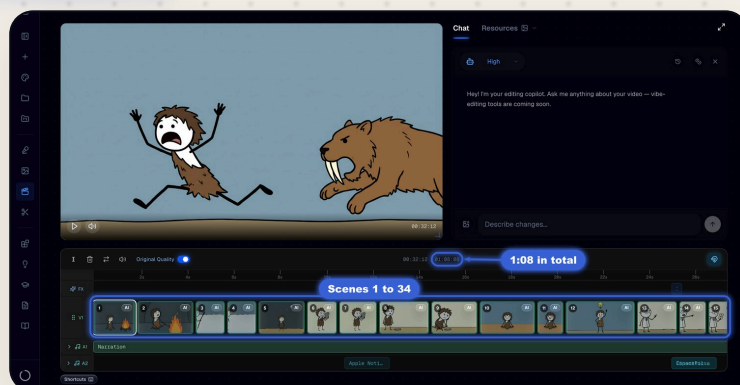


LEFT: OUR STYLE MASTER • RIGHT: THE ORIGINAL AT 0:01, 0:05, 0:26

6 • PASTE THE NARRATION AND GENERATE

For a fair test we used the original narration: the first 59 seconds, **177 words**, taken from the audio with speech to text, not from YouTube's captions, which add tags such as [music]. A script needs 900 characters or more. For a video you will publish, write your own.

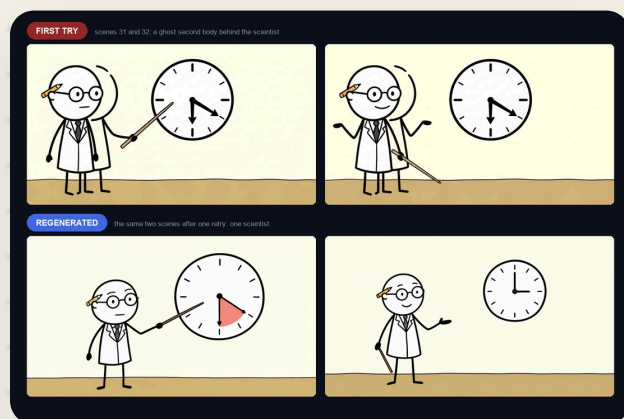
In the **Creation Lab** pick your template, click **Script**, paste the narration and set **Visuals Quality** to **Extra High**. We left animation off, because the original does not move. It finished in **9 min 42 s: 34 scenes in 68 seconds**, with a narrator voice and 5 sound effects.



THE FINISHED JOB • THE TIMELINE HOLDS ALL 34 SCENES, 1:08

7 • LOOK AT EVERY SCENE AND FIX THE BAD ONES

Open the editor and look at each scene at full size. We checked all 34. The caveman keeps the same design in the 22 scenes where he appears, and the scientist in his 12. Two were wrong: in scenes 31 and 32 the scientist had a ghost second body. We regenerated those two in one batch, 104 seconds, and both came out with one scientist.



TOP: FIRST TRY • BOTTOM: AFTER ONE RETRY



SCENE 31 IN THE EDITOR AFTER THE FIX

THE RESULT

The original: 30 pictures in 59.2 s. Ours: 34 pictures in 68.3 s, because the narrator speaks 15% slower (155 words per minute against 179). The original has 8 text cards such as 40,000 YEARS AGO and 6 HOURS. Ours has none: add your own in your editor. The side by side of six moments of both videos is on the web page of this guide.

THE 7 REFERENCES, READY TO COPY

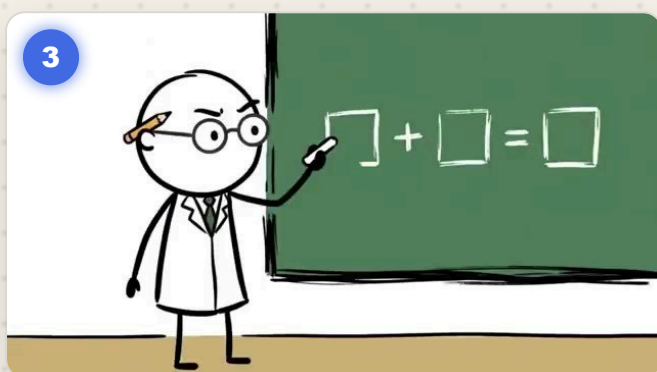
Seven frames from the first minute of the original, numbered 1 to 7. Save them, attach them to ChatGPT or add them to a template, in this order.



1 Character proportions, simple creature, flat snow



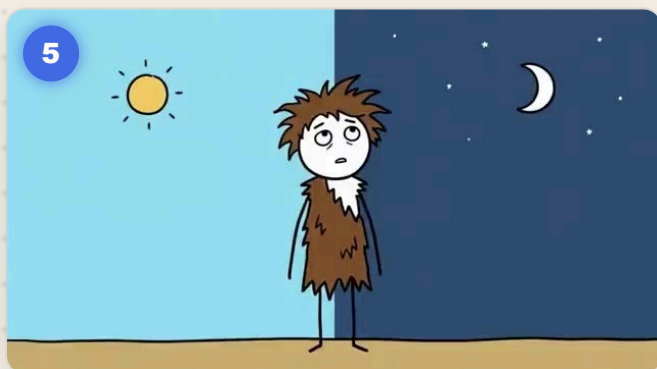
2 Expressive pose, prop symbol, cream background



3 Stick figure scientist variant and prop styling



4 Savanna environment, horizontal split, tracking



5 Day and night flat split background palette



6 Cave interior, flat campfire glow, seated pose



7 Action pose, stylized boar, clean chart graphic

SAVE THEM ON YOUR PHONE

Open the web page of this guide, press and hold each image to save it, or save the sheet with all 7 together and attach it to ChatGPT in one go.

calliopelabs.co/guides/stick-figure-history-video

COPY THESE PROMPTS AS THEY ARE

They describe the look of one video, so change the character and the places for yours. Take your own frames from the part you want to copy and label each one for what it shows.

VISUAL STYLE

HOW EVERY IMAGE LOOKS · COPY AS IS

PROMPT TEMPLATES (USE THESE FORMATS)

A. Character Scene Template

Minimalist 2D stick-figure caveman with a white circular head, messy brown hair, dot eyes, ragged brown fur tunic, and stick limbs, [INSERT ACTION OR POSE], [INSERT EXPRESSION], flat [INSERT ENVIRONMENT].

B. Object/Location Template

Simple hand-drawn 2D [INSERT OBJECT OR PROP] with thick black marker outlines, solid flat color fill, centered on flat [INSERT ENVIRONMENT].

GENERAL IMAGE STYLE (FOLLOW THIS CONSTRAINTS)

JSON

```
{
  "channel_visual_style": {
    "medium": "2D digital stick-figure illustration,
whiteboard cartoon style.",
    "aesthetic": "Minimalist educational webcomic, flat
and graphic.",
    "palette": "Cream white, warm tan, muted sky blue,
```

```
dark slate grey, fur brown, accent red and yellow.",
    "lighting": "Completely flat lighting; no directional
light, drop shadows, or gradients.",
    "composition": "Centered eye-level framing, massive
negative space, low horizon baseline.",
    "subject_treatment": "Thick black stick limbs, bold
white circular heads, simple expressive line faces, flat-
colored tunics.",
    "environment_treatment": "Extremely sparse horizontal
bi-color bands or single flat-color backgrounds.",
    "render_notes": "Uniform thick black marker outlines,
flat cell fills, zero textures or 3D depth."
  }
}
```

OTHER CONSTRAINTS

No 3D elements, CGI, or photorealism.
 No gradients, soft shading, or lighting effects.
 No textured brushes or complex backgrounds.
 No detailed anatomy or facial shading.
 Never generate watermarks, brand logos, or borders.

VIDEO STYLE

HOW THE CLIPS MOVE · COPY AS IS

No rules for video motion. Animate as desired and/or specified in other instructions.

REFERENCE LABELS

THE 7 FRAMES, IN THIS ORDER

- 1 Character proportions, simple creature, flat snow
- 2 Expressive pose, prop symbol, cream background
- 3 Stick figure scientist variant and prop styling
- 4 Savanna environment, horizontal split, tracking
- 5 Day and night flat split background palette
- 6 Cave interior, flat campfire glow, seated pose
- 7 Action pose, stylized boar, clean chart graphic

STYLE MASTER PROMPT

FOR STEP 5 · ON THE WEB PAGE

The prompt that builds the Style Master sheet is too long for this page. Copy it from the web page of this guide: calliopelabs.co/guides/stick-figure-history-video

MAKE YOUR OWN

Same steps for any video: check that it is an outlier, copy the look, test it, generate and look at every scene.

calliopelabs.co