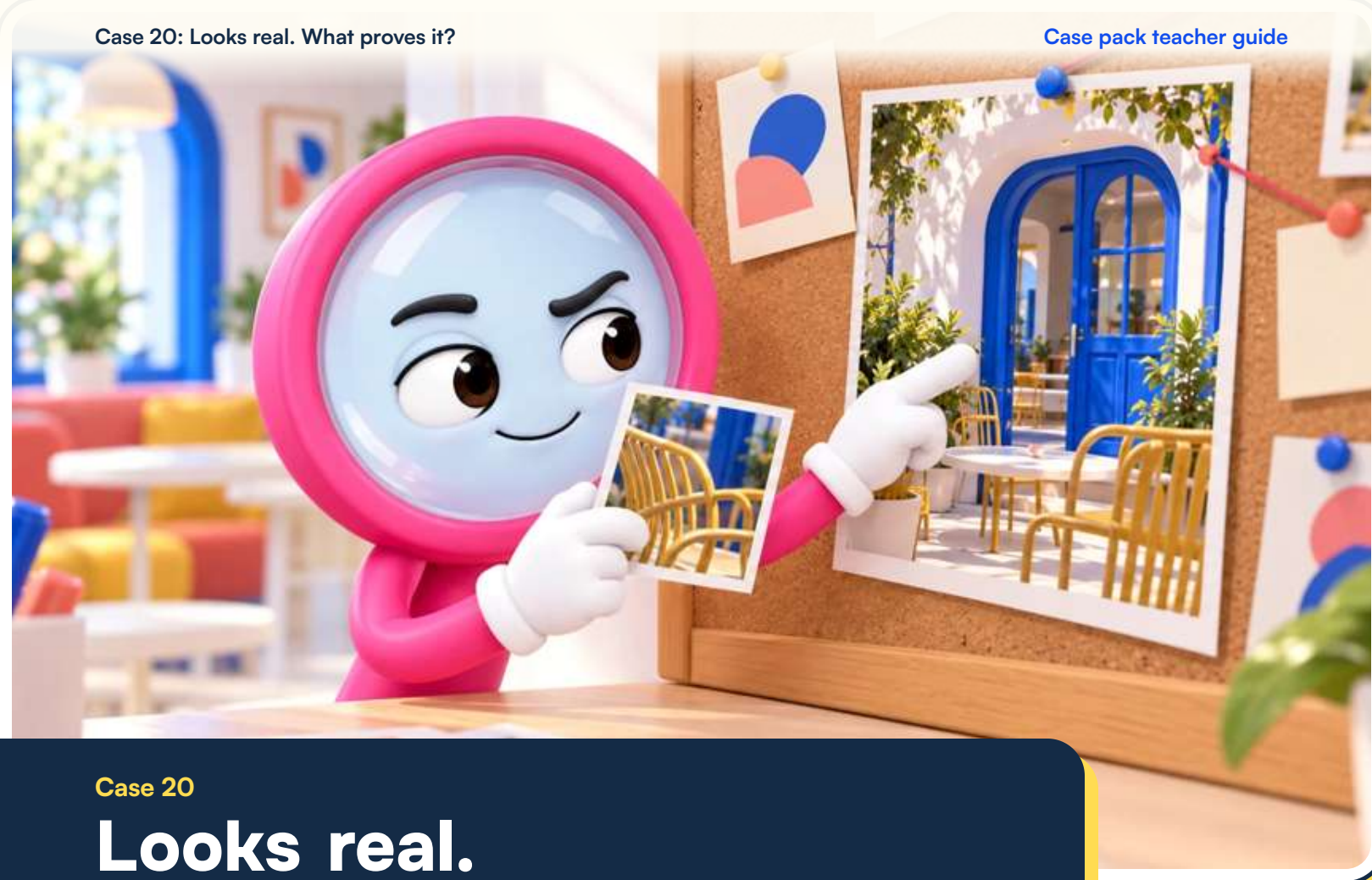




Case 20

Looks real. What proves it?

A realistic image says the courtyard flooded today. What would it take to post that as true?

Generated and recycled images now move through group chats as fast as real ones, and organizations that repost them lose trust quickly. Students who can write an honest verification note, including “we can’t confirm this yet,” protect their audience and their own credibility.

In the real world

NPR, 2023 Fake viral images of an explosion at the Pentagon were probably created by AI

CNN, 2023 ‘Verified’ Twitter accounts share fake image of ‘explosion’ near Pentagon, causing confusion

Jordy

I wanted a clean yes or no for the recap. “We can’t verify it yet” is less fun. It’s also what’s true.

You’ll make

Note (50—80 words) plus 120-word rationale

Time

About 50 minutes

Level

Grades 9—10



The case file

Look at what's actually here before deciding what it means.

Fictional case material

1

Courtyard flood image with caption, shared Wednesday 2:48 p.m.

Image, described: a photo-like view of the campus courtyard with water over the paving and benches; a person in the foreground holds up a phone.

Caption: "Courtyard flooded today, stay away." No link, no date on the image and no original poster named.



Where did this image first appear, and when? What does the caption claim that the image alone cannot show?

Fictional case material

2

Sami's zoom-in of the hand, crew chat, Wednesday 3:02 p.m.

Zoomed screenshot of the foreground hand: the fingers seem to merge at the knuckles.

Sami: "The hand is weird. But who posted this first? That's what I want to know."



What can a strange detail tell you, and what can't it?

Fictional case material

3

Reply in the crew chat, Wednesday 3:05 p.m.

"My roommate's friend said she saw water near the library earlier."



Is this corroboration? Who actually saw what, where and when?

Fictional case material

4

Result from a free online image detector (illustrative)

"12% likely AI-generated."



What can a score tell you, and what can't it? Does it say anything about whether the courtyard flooded today?

Outside the story

The case is fictional. The questions in it are not. Here is where they show up in the real world.

Still missing

Who first posted the image and when, whether it shows this courtyard on this day, whether anyone with a dated photo saw water, and how quickly the facilities page would report a same-day flood.

Real-world evidence



Real source

National Institute of Standards and Technology, Reducing Risks Posed by Synthetic Content: An Overview of Technical Approaches to Digital Content Transparency

Summary, not a quotation: NIST describes approaches such as provenance data, labeling, watermarking and detection for synthetic content, and explains that each has technical limitations, so no single approach settles on its own whether content is synthetic.



article

Fake viral images of an explosion at the Pentagon were probably created by AI [NPR, 2023-05-22](#)

A realistic fake image spread fast and briefly moved markets before officials confirmed nothing happened.



article

'Verified' Twitter accounts share fake image of 'explosion' near Pentagon, causing confusion [CNN, 2023-05-22](#)

A blue check was not proof: accounts that looked trustworthy still shared an image with no real source.

Discuss before deciding

- Which item did you trust first, and why?
- What would you need to see before repeating the strongest claim?
- Who is missing from these documents?

What the evidence can hold up

Two claims are in play: the flood and the image. Sort each piece of evidence under the claim it actually bears on. What would settle either one?

Evidence	Claim 1 The courtyard flooded today.	Claim 2 The image was generated by AI.
1 The image shows water in a courtyard, with no date or source	Supports	—
2 Facilities alerts page at 3:10 p.m.: no active alerts	Challenges	—
3 Secondhand report of water near the library	Supports	—
4 An odd hand and a 12% detector score	—	Leaves open

Still open

- The image’s earliest source and date

Words that do the work

Provenance

Procedencia

Information about where a file came from and how it was handled.

Sami wants to know who first shared the courtyard image, not just how the hand looks.

Corroboration

Corroboración

Independent evidence that supports a claim from another direction.

A campus facilities alert or another person’s dated photo would corroborate a flood. The image alone can’t.

Verification note

Nota de verificación

A short public note saying what was checked, what was found and what remains unknown.

Jordy decides the recap gets a verification note instead of the image.

Burden of proof

Carga de la prueba

The idea that whoever makes a claim should supply the evidence for it.

The caption claims a flood today. It’s on the claim, not on Jordy, to show where and when.

One student's draft

Write the verification note Jordy posts instead of the image, plus a short explanation of your choices.

Verification note: We've seen the courtyard flood picture going around. We checked the campus facilities alerts page at 3:10 and it shows no flooding today¹, and we couldn't find who posted the image first². The picture is fake, so please ignore it.³ Check the official alerts page before you change plans. Why I wrote it this way: I didn't repost the image because people would share it without reading the note. I listed what we checked so followers can check it too. The hand in the picture also looks weird, which shows it was made with AI⁴.

- 1** Strong: a named source and a time, so followers can check it themselves.
- 2** Honest about a gap in provenance, which is part of what was checked.
- 3** Revise: the evidence shows the flood is unconfirmed, not that the image is fake. It could be a real photo from another day or place.
- 4** Revise: an odd detail is a reason to look closer, not proof. Visual clues and detectors both have error rates.

Works well Uses evidence Gap to fix Think about the reader

One revision

Before

The picture is fake, so please ignore it.

After

We can't confirm the flood, so we're not sharing the image as news, and we'll update here by 6 p.m.

The revision reports exactly what the evidence supports and gives followers a time to check back, instead of a verdict the crew can't defend.

Teacher commentary

The note does the hard part well: it names a source, a time and a provenance gap, and it keeps the image out of the feed. It then overclaims twice, calling the image fake and treating an odd hand as proof of AI. Replacing both verdicts with calibrated language moves the response from Developing to Proficient on Reasoning.

Teach it in one period

Grades 9—10

Write the verification note Jordy posts instead of the image, plus a one-paragraph explanation of your choices.

Sources

All case materials and the NIST summary.

Objective

Students write a public verification note that reports what was checked without claiming a verdict the evidence cannot support, and justify their hedging.

Product

Note (50—80 words) plus 120-word rationale

Assessed by

The note is scored with the case rubric, and the rationale is checked for an explanation of why “can’t confirm” is more accurate than “fake.”

45 minutes	You	Students	Check for
5 min The image	Describe the courtyard image and caption: “Courtyard flooded today, stay away.” No link.	Rate belief 1—5 and write what makes it convincing.	Reasons are about appearance, not source.
8 min Zoom-in trap	Show the class zooming into a hand. Ask: what does an odd hand prove?	Sort clues into “reason to look closer” and “proof.”	Students put the hand under “reason to look closer.”
12 min Trace it	Hand out the earliest post found, the campus facilities page and a weather note.	Build a chain: claim, earliest source, date, corroboration.	Chains show no independent record of a flood today.
10 min Read the limits	Summarize NIST: provenance, labels, watermarks and detection each have gaps.	Write what a detector score can and can’t tell you.	Answer mentions error rates or missing provenance.
7 min Write the note	Model: checked / found / still open.	Draft a 50—80 word verification note for the recap.	Note states uncertainty without calling anyone a liar.
3 min Exit	Ask: what would change your confidence?	Name one piece of evidence.	Answer names independent, dated evidence.

Before class

Prepare the case documents, one set per pair:

- ✓ Courtyard flood image with caption, shared Wednesday 2:48 p.m.
 - ✓ Sami’s zoom-in of the hand, crew chat, Wednesday 3:02 p.m.
 - ✓ Reply in the crew chat, Wednesday 3:05 p.m.
 - ✓ Result from a free online image detector (illustrative)
- Every document is fictional case material, printable from the online field guide.

Questions to keep asking



Lupa
No link, no date, no original poster. The image arrived without its backstory.



Duda
What would you expect to find if the courtyard really flooded today?



Pauta
A detector score is a clue with error rates. It isn’t a finding.

Support every student, then assess

Success looks like

- ✓ States what was checked
- ✓ Avoids a verdict the evidence can't support
- ✓ Gives followers a next step
- ✓ Rationale explains hedging choices

Supports at this level

- Note template: checked / found / open
- Model note from another topic

Access

- Use a printed image-check card with icons
- Describe the image aloud for students with low vision
- Spanish vocabulary: imagen, fuente, verificado, sin confirmar
- Offer an oral verification note
- Reduce to two checks: who posted first, is there another record

Stretch

- Compare how two news organizations describe their image-verification process
- Research what content credentials record and what they cannot show

Challenge: Explain why “we can’t confirm” is more accurate than “fake.”

Rubric	Beginning	Developing	Proficient	Advanced
Evidence	Treats the image as proof.	Cites a detector or a detail as a verdict.	Checks the earliest source and an independent record.	Checks several sources and names what each can't show.
Reasoning	Declares real or fake with no basis.	Hedges without explaining why.	Matches confidence to evidence.	Separates authenticity of the file from truth of the event.
Audience & purpose	Urges spreading.	Warns without guidance.	Gives followers a safe next step.	Gives a next step, a follow-up time and a way to contribute evidence.
Revision	No change.	Softens words only.	Replaces the verdict with a verification note.	Explains how each change protects followers and credibility.

Standards this can show

SL.9-10.2 Common Core 9–10

RI.9-10.8 Common Core 9–10

W.9-10.8 Common Core 9–10

RI.11-12.7 Common Core 11–12

W.11-12.8 Common Core 11–12

SL.11-12.3 Common Core 11–12

Grades 6—8

With a partner, fill an image-check card and write two sentences telling friends whether to trust the courtyard picture.

Use

The image description, caption and campus facilities page only.

You'll make

Completed card plus 2—3 sentences

Show it your way
Write it

Write 2—3 sentences with the frame “The picture shows __. It doesn’t show __.”

Say it

Tell a partner in 30 seconds whether you would share the picture and which proof you would want first.

Show it

Draw the image-check card as a four-box chart and circle the box that is still empty.

Done when

- ☐ Separates what the image shows from what it claims
- ☐ Names a better source to check
- ☐ Uses “unverified” correctly
- ☐ Doesn’t accuse the poster

Item 1: what it shows, and what it doesn’t**Item 2: what it shows, and what it doesn’t****Item 3: what it shows, and what it doesn’t****What I can actually claim****Need a boost?**

- Image-check card: claim, who posted first, date, other proof
- Frames: “The picture shows __. It doesn’t show __.”
- Vocabulary cards: source, evidence, unverified
- Talk before writing



Prefer typing? Scan to do this in the online workspace.

Grades 9—10

Write the verification note Jordy posts instead of the image, plus a one-paragraph explanation of your choices.

Use

All case materials and the NIST summary.

You'll make

Note (50—80 words) plus 120-word rationale

To A real role

How it gets there Email, comment,
letter, DM

Opening Who you are and why you're writing

The ask One clear request

Your evidence Two specifics beat five adjectives

Close What happens next

Done when

- ☐ States what was checked
- ☐ Avoids a verdict the evidence can't support
- ☐ Gives followers a next step
- ☐ Rationale explains hedging choices

Need a boost?

- Note template: checked / found / open
- Model note from another topic



Prefer typing? Scan to do this in the online workspace.

Honors

Argue whether platforms should be required to label AI-generated images.

Use

NIST overview plus one teacher-provided article on labeling policies.

You'll make

Argument, 550—650 words

Thesis A claim someone could reasonably disagree with

Done when

- ☐ Clear, qualified position
- ☐ Uses technical limits accurately
- ☐ Addresses counterarguments
- ☐ Proposes a realistic policy

Reason 1: claim, evidence, commentary

Reason 2: claim, evidence, commentary

Need a boost?

- Claim/counterclaim chart
- Examples of labels

The strongest counterargument, and your answer

Why it matters beyond this case



Prefer typing? Scan to do this in the online workspace.

Pre-AP

Closely observe the courtyard image and its caption, record what can and can't be seen, and write an evidence-based verification paragraph.

Use

The image description, caption and campus alerts page.

You'll make

Observation grid plus paragraph (150–200 words)

Item 1: what it shows, and what it doesn't**Item 2: what it shows, and what it doesn't****Item 3: what it shows, and what it doesn't****What I can actually claim****Done when**

- ☐ Observations are separate from claims
- ☐ Confidence matches evidence
- ☐ Cites the campus alerts page
- ☐ Revision reflects discussion

Need a boost?

- Observation grid: in the image / claimed by the caption
- Question stems: "What would I expect to see if this were today?"
- Paragraph frame



Prefer typing? Scan to do this in the online workspace.

AP English Language

Analyze how the viral caption persuades, then write a verification note that uses the same brevity honestly.**Use**

The caption, the image description and two reposts with added comments.

You'll make

Analysis (about 400 words) plus note

Thesis A claim someone could reasonably disagree with

Done when

- ☐ Thesis on how the caption persuades
- ☐ Analyzes specific word choices
- ☐ Note applies the analysis
- ☐ Commentary explains effect

Reason 1: claim, evidence, commentary

Reason 2: claim, evidence, commentary

Need a boost?

- Rhetorical situation grid
- Commentary stems on urgency and imperatives

The strongest counterargument, and your answer

Why it matters beyond this case



Prefer typing? Scan to do this in the online workspace.

AP U.S. Government

Argument essay: should government regulate AI-generated images of public events, or leave it to platforms and users?

Use

Case, NIST overview, and a short explainer on First Amendment limits on regulating speech.

You'll make

Argument essay, 40 minutes

Thesis A claim someone could reasonably disagree with

Reason 1: claim, evidence, commentary

Reason 2: claim, evidence, commentary

Done when

- ☐ Defensible claim
- ☐ Accurate First Amendment reasoning
- ☐ Uses case and NIST as evidence
- ☐ Responds to opposing view

Need a boost?

- Argument outline
- Vocabulary: prior restraint, content regulation, media literacy

The strongest counterargument, and your answer

Why it matters beyond this case



Prefer typing? Scan to do this in the online workspace.

AP Seminar

Individual written argument on “Can we still trust what we see?”, using a stimulus packet.

Use

Stimulus: NIST overview, Adobe’s labeled examples, one journalism verification guide and one psychology article on belief.

You’ll make

Individual written argument, 1,200—1,500 words

Thesis A claim someone could reasonably disagree with

Done when

- ☐ Clear research question
- ☐ Evaluates technical, journalistic and psychological perspectives
- ☐ Original synthesis
- ☐ Feasible norm proposed

Reason 1: claim, evidence, commentary

Reason 2: claim, evidence, commentary

Need a boost?

- Stimulus annotation
- Perspective map
- Argument outline

The strongest counterargument, and your answer

Why it matters beyond this case



Prefer typing? Scan to do this in the online workspace.

Pre-IB

Trace the courtyard image as far as the case allows and write a verification analysis: what each source can establish, what it cannot and how confident the crew should be. Then write a 300-word reflection on the question: when is seeing still believing?

Use

The image description and caption, Sami's zoom-in, the detector result, the facilities alerts page, the secondhand report and the NIST overview.

You'll make

Verification analysis (400—500 words) plus a 300-word reflection

When**What happened, and how we know**

When**What happened, and how we know**

When**What happened, and how we know**

When**What happened, and how we know**

When**What happened, and how we know**

The turning point, and the missing step

Done when

- ☐ States what each source can and cannot establish
- ☐ Keeps the authenticity question apart from the event question
- ☐ Uses calibrated language for confidence
- ☐ Reflection develops a claim and a counterclaim about visual evidence

Need a boost?

- Source chain template: claim, earliest source, date, corroboration
- Confidence scale with wording for each level
- Reflection stems: "I can know by looking when..., but not when..."



Prefer typing? Scan to do this in the online workspace.

IB MYP

Investigate “Can we still believe our eyes?” and make a short guide for younger students.

Use

Case materials and Adobe’s labeled examples; global context: Scientific and technical innovation.

You’ll make

Guide plus 150-word reflection

Frame 1



What we see and hear

Frame 2



What we see and hear

Frame 3



What we see and hear

Frame 4



What we see and hear

Done when

- ☐ Statement of inquiry links concepts
- ☐ Accurate verification steps
- ☐ Practice example is labeled
- ☐ Reflection connects to the global context

Need a boost?

- Statement of inquiry frame
- Guide template
- Source checklist

Why this order works What does the viewer know by the last frame that they didn’t at the first?



Prefer typing? Scan to do this in the online workspace.

IB DP Language and Literature

Write a guided textual analysis of a teacher-selected, clearly labeled AI-generated image post and its caption, answering the guiding question, then test your analysis against the case's courtyard post and its reposts.

Use

One teacher-selected image post that is clearly labeled as generated, plus the case's courtyard caption and two reposts with added comments.

You'll make

Guided analysis, 600—800 words, written in 75 minutes

Thesis A claim someone could reasonably disagree with

Reason 1: claim, evidence, commentary

Reason 2: claim, evidence, commentary

The strongest counterargument, and your answer

Why it matters beyond this case

Done when

- ☐ Responds directly to the guiding question
- ☐ Analyzes image and caption together with precise references
- ☐ Explains how the text positions its reader to act
- ☐ Sustains an organized line of analysis in clear academic language

Need a boost?

- Guiding question printed above the text
- Annotation layers: composition, realism cues, caption syntax, imperatives
- Glossary of visual terms: framing, foreground, focal point



Prefer typing? Scan to do this in the online workspace.