

Grading That Holds Up

Validity, accommodations, privacy and reliability — the parts of ASL assessment that carry real risk for students, and real exposure for you.

A grade is a claim about what a student can do. Most ASL grading problems are not strictness problems.

They are validity problems.

Effectiveness, Not Resemblance

Start here. This one can harm a student.

A rubric row like “handshape accuracy” or “clear production” will penalise a student with a tremor, limb difference, arthritis, cerebral palsy, hypermobility or motor-planning differences — for something that has nothing to do with language competence. Deaf signers with these conditions are fully fluent. Their language is intact; their production simply is not identical to a citation form.

Score what a signer would actually notice. Every row below is the more valid construct for *every* student, not only the ones with a documented difference — which also means no student has to disclose anything for your rubric to be fair.

INSTEAD OF

“Handshape accuracy.”

Was the handshape distinguishable from its minimal pairs?

Meaning-based, and it is the question a conversation partner is actually asking.

INSTEAD OF

“Clear, fluid production.”

Could a signer follow this at conversational speed without asking for repair?

Fluency is a comprehension outcome, not an aesthetic one.

INSTEAD OF

“Correct use of signing space.”

Were referents placed consistently enough to track?

Small space is only a problem if it costs the listener the referent.

INSTEAD OF

“Matches the model.”

Was the meaning clear to a signer?

Resemblance to a citation form is not the construct. Being understood is.

Two Kinds of Accommodation — *One preserves the measurement. One quietly replaces it.*

Both columns are legitimate supports. The difference is what the resulting score still means — so if you need something from the right-hand column, say so on the rubric and score the remaining criteria rather than pretending the target is unchanged.

KEEPS THE TARGET INTACT

Still measures ASL

- Extended time on a receptive task
- The passage chunked into shorter segments
- A slower signing rate
- Repeated viewings
- A live conference instead of a recording
- A scribe for written portions
- Scheduled breaks, or assessment across two sittings
- Seating, lighting, or camera position changes

CHANGES WHAT YOU MEASURE

Now measuring something else

- English text of the signed passage — this becomes a reading test
- A written transcript in place of demonstrated comprehension
- Captions on a receptive assessment
- Spoken English responses on an expressive task
- Grading gloss for spelling and formatting
- Dropping the interpersonal component entirely

RULE OF THUMB Ask what the score would still let you claim. If a support changes the answer to “**can this student understand ASL?**” into “**can this student read English?**”, the construct has moved and the grade no longer means what your gradebook says it means.

Don't Grade English by Accident The most common silent validity leak.

INSTEAD OF

A Deaf culture unit assessed by English essay.

Signed response, multiple choice, sequencing, or diagramming. An essay measures English writing alongside the content, and the two scores are inseparable afterwards.

INSTEAD OF

Receptive comprehension answered in written English.

Answers in ASL, drawing, matching, or ordering. Otherwise a strong signer with weak writing scores as a weak signer.

INSTEAD OF

Marking gloss for accuracy.

Accept gloss with generous tolerance for spelling and formatting. Glossing is a notation system, not the language. Grading it as though it were is contested at best.

Grade Against the Level Not against fluency.

LEVELS-AREN'T-PERCENTAGES

Novice-Mid is not seventy percent of a Deaf adult. Proficiency levels and percentage scales are two different measurement systems. Averaging one into the other produces a number that means nothing.

ON-LEVEL-IS-AN-A

A Level 1 student performing exactly at Level 1 expectations has earned an A. If that feels generous, the discomfort is with the level description, not the student.

CHECK-YOUR-TOP-BAND

If the top band describes near-native production, every beginner is capped at mediocre by design. Write the top band as the ceiling for *this* course, not for the language.

RECENCY-OVER-AVERAGING

Week fourteen is better evidence than the mean of weeks one through fourteen. Averaging a growth curve punishes students for having once been beginners.

Questions to Take Upstairs _____ *For your administrator and technology director.*

Once you collect video of minors signing, you are holding student records with their faces in them. Work through these before the next unit rather than after an incident.

- ☐ Where is student video stored, for how long, and who else can access it?
- ☐ Do the platform’s terms permit it to retain uploaded content, or to train models on it?
- ☐ Does peer feedback expose one student’s record to another — and what policy permits that?
- ☐ What is the consent process, and what happens when a family declines?
- ☐ Is there a deletion date, and does anyone actually run it?
- ☐ Who owns the recording if a student transfers out, or asks for it to be removed?

STUDENTS WHO MAY NOT BE ABLE TO SUBMIT VIDEO AT ALL

▪ No private space to record in	▪ No reliable device or data
▪ A family that will not consent	▪ Foster placement or a custody arrangement
▪ Safety or immigration concerns	▪ Religious modesty
▪ Gender dysphoria	▪ A trauma history

So: keep an equivalent live or in-person path open at all times, and never require a student to explain why they need it. If video is the only route to a grade, these students are locked out of your gradebook.

NOT LEGAL ADVICE

District policy and state student-privacy law vary, and this page is not a legal opinion. Treat it as a list of questions for the people in your building whose job it is to answer them — and get the answer in writing where you can.

Norming With Anchors

The highest-leverage thing a department can do together.

- 1

Collect two or three student performances at each rubric level, with permission and identities handled properly.
- 2

Everyone scores them independently, with no discussion first.
- 3

Reveal the scores and argue until you agree where each one sits.
- 4

Keep the agreed clips. They are now your anchor set.
- 5

Recalibrate once a semester, and whenever a new teacher joins the department.
- 6

When a grade is challenged, the anchor set is your evidence that a 3 means the same thing in every section.

ONE-CRITERION-AT-A-TIME	Grade one criterion across all students, then the next. Cuts rater drift, and it is faster than working student by student.
SHUFFLE-THE-ORDER	Change the order you grade in. The first five and the last five performances get systematically different treatment.
NAME-THE-HALO	The outgoing, likeable student reads as more fluent than they are. Knowing the effect has a name is most of the defence against it.

Integrity by Design

Detection is unreliable. Design is not.

Re-recording until perfect, a sibling or CODA friend signing off camera, and now AI-generated signing avatars and filters. Accusing a student you cannot prove anything about does more damage than the cheating did — so remove the opportunity instead of hunting for it.

PROMPT-AT-THE-MOMENT	Give the prompt when the assessment starts, not before. Most of the workaround economy depends on knowing the task in advance.
MAKE-IT-LIVE	A conversation cannot be pre-recorded. One unrehearsed two-way component does more than any detection tool.
WEIGHT-THE-ROOM	Gather most evidence in class. When the bulk of the grade comes from the room, the incentive to fake the rest collapses.

What Protects You

If you stop grading everything, someone will ask about rigour.

EVIDENCE-NOT-SCORES

Keep lightweight records of formative work: a roster with checkmarks, dated progress notes, a folder of ungraded video. “I don’t score practice” should look like a design decision, not an absence.

SAY-IT-IN-THE-SYLLABUS

One paragraph on what is graded and what isn’t. Written down in advance, it is policy. Explained afterwards, it sounds like a defence.

SAY-IT-OUT-LOUD

Explain it at curriculum night. Before a parent asks why there is no homework grade, not after.

Three More

Small, and easy to get wrong.

NO-ZEROS-ON-100

A zero on a hundred-point scale is mathematically catastrophic. It is almost never the consequence anyone actually intended. Use the four-point scale, or a floor.

COMMUNITY-IS-NOT-A-RESOURCE

Deaf community assignments carry access and burden costs. Transport, money and geography limit who can go — and thirty students descending on one Deaf coffee chat is a cost to that community. Build alternatives in from the start, and ask before you send.

KNOWLEDGE-NOT-STANCE

Assess what a student knows, not which side they take. Deaf culture content slides easily into grading whether a student holds the “right” view on questions still contested inside the community itself.

The short version: make sure the grade is about the language, make sure every student can reach the assessment, and make sure your department means the same thing by a 3. Rigour that measures the wrong thing is not rigour.

RIGHT THING
MEASURED RIGHT