

What this pilot shows. In one school day, the Mission Analysis System recorded structured behavioral evidence on about 60 students during a single mission. The data names specific coaching targets and shows which of the practices the mission was built to elicit showed up on the bridge, and which did not. It is a formative read for coaching and planning, not a summative measure of learning.

01 / BY THE NUMBERS

What the day produced

124

SEVENTH GRADERS COMPLETED THE FULL MISSION

60

STUDENTS BEHAVIORALLY MEASURED BY THE SYSTEM

3

SPECIFIC COACHING TARGETS IDENTIFIED

02 / STANDARDS MAPPING

What the mission was built to elicit

The mission was engineered to elicit 18 TEKS Student Expectations across four subjects, and each of the 10 measured sessions was scored for evidence of each one. The rows below name the mapping as what it is. The science expectations are the practices and recurring themes of §112.27(b), the habits of scientific work, not grade 7 science content.

SUBJECT	CITATION	WHAT THE MAPPED ROWS ARE	SES
SCIENCE	§112.27(b)(1)–(3)	Scientific and engineering practices: asking questions, analyzing data, developing and communicating explanations, arguing from evidence	7
SCIENCE	§112.27(b)(5)	Recurring themes and concepts: patterns, cause and effect, systems and system models, stability and change	4
ELA	§110.23(b)(1)	Oral language: listening actively, following and giving oral instructions, engaging in discourse	3
MATH	§111.27(b)(4)	Proportionality: unit rates, and problems involving ratios, rates, and percents	2
TECH APPS	§126.18(c)	Technology terminology in team collaboration, and collaborating on a digital platform	2

NO CONTENT PAIRING WAS RECORDED FOR THIS DEPLOYMENT.

What this means. Five of the mapped expectations reached Medium evidence in 4 or fewer of the 10 measured sessions, including active listening and digital collaboration. That is useful information. It tells us where the mission design needs adjustment for the next deployment, and it tells classroom teachers where targeted instruction would have the highest leverage. One of the two Technology Applications expectations describes collaborating on a digital platform, which a voice only bridge mission does not ask for. Thin evidence there is a mission design fact, not a student gap.

How evidence is scored. Each measured session is scored for evidence of each mapped expectation by rules fixed before the mission runs. It is not a grade, and thin evidence abstains rather than reporting a zero. This edition reports no subject percentages.

About this sample. This is a de-identified edition of a delivered pilot brief. Campus, district, and staff names are withheld. Session data and limits are as delivered. This revision labels the standards mapping more precisely than the delivered edition and drops its subject coverage percentages.

FINDINGS & LIMITS

WHAT THE DATA SHOWS — AND WHAT IT DOESN'T

03 / WHAT THE DATA SHOWED

Student behavior patterns

STRENGTH

SCIENTIFIC ARGUMENTATION

10 of 10 measured crews produced strong evidence of engaging respectfully in scientific argumentation using applied scientific explanations and empirical evidence (TEKS §112.27(b)(3)(C)). The mission format makes this happen naturally.

COACHING OPPORTUNITY

ACTIVE LISTENING

Active listening and clarifying questions (TEKS §110.23(b)(1)(A)) reached Medium evidence in 4 of 10 crews. This is the most direct coaching target for follow up classroom work and a natural entry point for Habit 5 on Leader in Me campuses.

SURPRISING PATTERN

ONE VOICE PER BRIDGE

Discourse Equity averaged 0.51 across the 10 measured crews. Each crew had a single dominant speaker. Building balanced participation is a clear, measurable goal for the next deployment.

ENGAGEMENT

52% SUSTAINED ENGAGEMENT

Just over half of mission time landed at or above each session's neural engagement baseline. Strong for a first exposure to a high stakes simulation. Whether repeat exposure lifts this number is a question a second deployment can answer. This pilot cannot.

04 / WHAT THIS PILOT DOES NOT SHOW

Honest limits

SINGLE DAY

NO GROWTH MEASUREMENT

One mission produces a snapshot of behavior. It does not measure whether students improved, retained skills, or transferred them to other settings. A second deployment with the same mission template would produce that comparison, and any comparison would be directional only.

NO COMPARISON GROUP

NO CAUSAL CLAIM

The data describes what students did during the mission. It does not prove the program caused those behaviors or that the same students would not have produced similar evidence in another setting. Causal claims require comparison data we do not yet have.

SAMPLE LIMIT

60 OF 124 STUDENTS

Behavioral measurement covers the 10 crews on Bridge A, which captured complete telemetry and audio. Bridge B had an audio capture fault and recorded nothing, so its 11 crews have no data. The host teacher caught the fault during the event. The root cause was identified afterward and remediated, and the fix was confirmed in later testing. A full coverage deployment would lift the sample to the entire grade.

FORMATIVE NUMBERS • NOT EXTERNALLY VALIDATED

Every score is computationally verified against telemetry and transcript inputs. None has been externally validated against trained human judgment. Treat the numbers as formative, a conversation starter for coaching, not a summative grade, an individual diagnosis, or a ranking against other schools.

NEXT STEPS

CLOSING A PILOT • USING THE DETAILED REPORT

05 / RECOMMENDED NEXT STEPS

Closing a pilot

01

30 MINUTE DEBRIEF

Walk through the data with the host teacher and an administrator. Identify which findings to share with teachers and which to use for next year's planning.

02

AGGREGATE DATA USE

Written confirmation that aggregate, de-identified data may appear in MSA case study materials. No student names. No individual data. Class wide patterns only. Naming the campus or the district requires separate written approval.

03

SECOND DEPLOYMENT

Run the same mission template with the same cohort or the next one. That comparison is what turns a baseline into a directional read on growth. Any comparison stays directional only until more cohorts run the same template.

06 / SUGGESTED READING PATH

How to use the detailed report

FOR THE PRINCIPAL • 5 MINUTES

PAGE 1 OF THIS BRIEF

The hero, the three numbers, the standards mapping. That is the one page a board update or a program review needs.

HOST AND SUBJECT TEACHERS • 15 MINUTES

PAGES 1 AND 2, THEN THE TEKS TABLES

Pages 1 and 2 of this brief, then the TEKS evidence tables on pages 7 and 8 of the Cohort Comparison Report. The five expectations with the thinnest evidence are the strongest candidates for follow up classroom work.

CURRICULUM AND PROCUREMENT • 30 MINUTES

THE FULL COHORT COMPARISON REPORT

Beginning with the methodology and limits block on page 2. The metric definitions, evidence definitions, and limits are what a procurement committee will want to verify before authorizing the program at scale. Every TEKS citation in both documents is quoted in full from the Texas Administrative Code.

THE COHORT COMPARISON REPORT RUNS EIGHT PAGES. BOTH DOCUMENTS DOWNLOAD TOGETHER.

This brief and the full eight page Cohort Comparison Report — metric definitions, crew by crew profiles, the full TEKS table with verbatim Student Expectations, and the privacy architecture your district will ask about — are both available from the sample report page. We walk them through in a 15 minute briefing, so your questions get answered in the room.

[DOWNLOAD](#) • [MOBILESPACEADVENTURES.ORG/REPORT](https://mobilespaceadventures.org/report) • [BOOK](#) • [/CONTACT](#) • [INFO@MOBILESPACEADVENTURES.ORG](mailto:info@mobilespaceadventures.org)