edtech impact

CROSS-VENDOR LEARNING REPORT · OCTOBER 2026

What schools want to know *before they trust edtech*

Data collected across four co-design workshops with school leaders, teachers, researchers and edtech industry experts shows what people evaluating edtech want to know, and when. Working with one vendor's product per workshop, groups wrote 189 questions and drafted 13 evaluation plans. Together they offer practical learning for edtech product evaluation, school and trust-led evaluations, and vendors preparing evidence for schools.

ABOUT THIS REPORT

- Post-its and worksheets from four co-design workshops at the EdTech Impact Summit, London, 8 October 2026
- 189 unique questions from 14 groups, and 13 evaluation worksheets
- One product per workshop: Tes Provision Map, LbQ (Learning by Questions), Magma Maths and Zero Gravity
- School leaders, teachers, researchers and edtech industry experts took part
- Presentation decks from LbQ and Zero Gravity, provided before the workshop
- Transcribed, coded and analysed by the EdTech Impact Testbed team

How the *workshops were run*

189
unique questions
from post-its written by 14 groups in 4 workshops

13
evaluation worksheets
sketching a 3 to 6 month testbed

40%
of short-term questions
were about fit, compliance and cost (34 of 84)

0 of 7
groups that answered
wanted the product company to draw conclusions from the data

Numbers are small (4 workshops, 14 groups, 13 worksheets), so read the figures in this report as indications, not statistics.

ABOUT EDTECH IMPACT AND THE SUMMIT

EdTech Impact is an independent edtech evaluation organisation: its always-on platform, Impact Manager, helps school leaders make faster, evidence-based decisions, and it independently collects, structures and presents evidence on edtech products. It designed and ran these workshops at the EdTech Impact Summit (London, 8 October 2026), the UK's first summit dedicated to edtech evidence in schools, with around 100 leaders asking how to use evidence to make smarter edtech decisions, cut wasted spend and free up teacher time.

The Summit ran four workshops, each focused on one product: Tes Provision Map, LbQ (Learning by Questions), Magma Maths and Zero Gravity. Each workshop lasted 75 minutes. Attendees (school leaders, teachers, researchers and edtech industry experts) took part across the four workshops: 14 in the LbQ workshop, 17 in Magma Maths, 16 in Tes and 16 in Zero Gravity (63 in total). Within each workshop, attendees formed 3 or 4 groups of about three to six people, and each group worked independently throughout. Attendees were asked to take the view of a MAT leader who has no time for a full sales demo every time staff ask for a new product, and who needs a template that lets their school or trust set up a testbed evaluation producing the evidence needed for a procurement decision.

Every workshop followed the same running order. The vendor gave a 10-minute presentation with a live demo of the platform, covering the intended age group, target users and context, and was briefed to end the 10 minutes with about three claims about impact. Each group then wrote post-its in answer to the prompt "What immediate questions come to mind regarding the potential product impacts that you would be interested to know the answers to?", one question per post-it (5 minutes). The prompt did not set a timeframe, but groups were encouraged to think of questions for the short, medium and long term. Once all its post-its were written, the group placed them on a timeframe sheet according to when each question could be answered: short term (immediate), medium term or long term (5 minutes).

Each group then filled in its own A1 worksheet (shown in Appendix A6) in 20 minutes, choosing a question suited to a 3 to 6 month testbed and planning the cohort, data points, data collection and analysis, and how positive, neutral and negative findings would be handled, followed by 5 minutes to write a research question using the structured prompt at the foot of the worksheet. Only after that did the vendor present a 5-minute slot on the evidence data it already holds and can share. This order was deliberate: vendors were asked to show mostly a demo and end with about three claims, mirroring what typically happens when a school leader first sees a product, such as a short pitch from an educator. The workshop therefore first captured the questions that first look raises, and then showed whether the vendor already holds evidence that answers them or whether schools are asking for evidence that does not yet exist. LbQ was a partial exception: because its evidence approach is part of how the product works, it covered some of its evidence at the start and added more in the evidence slot. Each workshop closed with every group reading out its research question; the planned discussion of where vendor data matched what schools wanted to measure did not happen in depth. Vendor staff in each workshop were allowed to answer questions while groups filled in the worksheet, but were asked not to influence the process, so that the learnings would not simply tell them what they wanted to hear.

The four workshops at a glance

Each workshop covered one product. Groups are counted within their workshop: Tes Group 1 and LbQ Group 1, for example, are different groups in different rooms. Attendee numbers are shown for context only; they are not used to calculate any figure in this report.

Workshop (product)	Attendees	Groups	Post-its kept	Worksheets
Tes Provision Map	16	3	25	3
LbQ (Learning by Questions)	14	4	83	4
Magma Maths	17	4	43	4
Zero Gravity	16	3	38	2
Total	63	14	189	13

Running order of each workshop

All four workshops followed the same 75-minute running order. Within a workshop, each group worked independently at its own table.

10 min
VENDOR

Presentation, mostly a live demo, ending with about three impact claims

5 min
EACH GROUP

Writes post-its, one question per post-it

5 min
EACH GROUP

Places its post-its on the timeframe sheet: short, medium or long term

20 min
EACH GROUP

Fills in its own worksheet (Appendix A6)

5 min
EACH GROUP

Writes a research question at the foot of the worksheet

5 min
VENDOR

Shows the evidence data it already holds

Close
ALL GROUPS

Each group reads out its research question

How to read the findings

- Timeframes (stages) are each group's own judgement of when a question could be answered, not the order in which they would ask it. All 18 post-its left unsorted ("Not assigned") came from LbQ groups.
- The post-it prompt asked for the immediate questions that came to mind after seeing the demo and hearing the claims, one per post-it. It did not set a timeframe but encouraged groups to think of questions for the short, medium and long term. Although it mentioned potential product impacts, many questions were about fit, features and set-up, which may itself be a finding.
- Despite the instruction to write questions, a few post-its are single words (for example "Outcome", "Engagement" and "Alignment" from Zero Gravity Group 1) or arrow lists (LbQ Group 3).
- The worksheet was completed in about 20 minutes, groups were steered towards a 3 to 6 month window, and the tick-box options were printed on the sheet. Worksheet answers are best read as starting instincts rather than worked-up plans.
- The LbQ workshop contributed the most data: its 4 groups wrote 44% of post-its (83 of 189), although it had the smallest room (14 attendees), and 4 of the 7 groups that answered the conclusions question were in the LbQ workshop. Each workshop covered only one product, so differences between products cannot be separated from differences in who was in each room, and answers from mixed groups are not split by role.

Summary of learnings for *school-led evidence building*

Each learning gives the share and count behind it (see Appendix A1). With 14 groups across four workshops, figures are indications, not statistics; LbQ workshop participants contributed the largest share of the data.

1 **The focus shifts across the timeframes.**

Sorted by when each question could be answered (not by importance), short-term questions were most often about fit, compliance and cost (40%, 34 of 84), medium-term ones about day-to-day use and teaching (64%, 34 of 53), and long-term ones about impact and evidence (41%, 14 of 34).

2 **Evidence of impact on pupil learning and outcomes was the most common theme.**

It covers 13% of all questions (25 of 189) and 26% of long-term ones (9 of 34). A few asked about the mechanism, what would have happened without the tool and who it works best for. The two vendor decks we have rely mainly on usage data, surveys and case studies (Section 3).

3 **Each product drew a different mix of questions.**

Tes drew integration and time saving, Magma Maths pedagogy, Zero Gravity data and reporting and pupil experience, and LbQ evidence of impact. With one product per workshop, this may also reflect who was in each room, but it suggests a common core plus a product-specific module (Section 4).

4 **Groups questioned how time savings are measured.**

Half of the 12 workload post-its (6) asked how a saving was quantified or baselined, for example whether it was "hypothetical or tested/piloted" (Tes Group 3).

5 **No group wanted the product company to draw the conclusions.**

None of the 7 groups that answered chose the product company to formulate conclusions from the data collected and analysed. 5 of 7 gave conclusions to the school or trust and 2 (both Magma Maths) to a researcher.

6 **Over half accepted the product company collecting data.**

5 of 9 groups ticked the product company to collect data, but only 2 of 8 gave it a share of the analysis, which sat mostly with internal teams (5 of 8).

7 **Attainment, teacher voice and pupil voice were listed most.**

Attainment appeared in 10 of 12 data-point lists, teacher voice in 9 and pupil voice in 8; usage data in only 4. Vendor evidence came after the worksheets by design, which may partly explain the lean towards assessment data.

8 **Most planned evaluations sat at the top of the 3 to 6 month window.**

7 of 13 chose about six months and 3 went beyond the window to a year. 6 of 13 relied on a pre/post baseline alone and 5 proposed a comparison group. Drafted in about 20 minutes, these are best read as starting instincts.

9 **Data protection and IT leads were widely included.**

7 of 9 groups ticked both the DPO and the IT lead for the evaluation team, and integration was the top short-term theme (15 of 84).

10 **Participants wanted pupils and parents involved.**

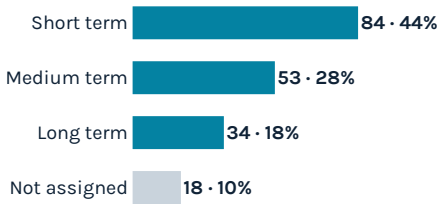
Pupils were ticked as a voice to include by 9 of 10 groups and parents by 7 of 10, and 4 worksheets also wrote pupils or parents into the evaluation team.

Questions asked moved from practical fit towards *evidence of impact*

Groups wrote 189 unique questions and sorted them into the timeframe in which each could be answered. Short-term questions most often asked whether the product would fit the school safely, medium-term questions about daily use and teaching, and long-term questions about evidence that it works. The table shows how many post-its, themes and groups each timeframe covered.

Timeframe	Post-its	Share	Themes covered	Groups	Top theme
Short term	84	44%	14 of 15	14 of 14	Integration, devices & tech fit (18%, 15 of 84)
Medium term	53	28%	13 of 15	13 of 14	Pedagogy & teaching practice (23%, 12 of 53)
Long term	34	18%	12 of 15	11 of 14	Evidence of impact & outcomes (26%, 9 of 34)
Not assigned	18	10%	8 of 15	3 of 14	Features & functionality (28%, 5 of 18)

Figure 1. Post-its by timeframe



Source: workshop post-its (189 unique questions), EdTech Impact Summit, October 2026

Figure 2. The focus shifts over time

	Short	Medium	Long	All
Impact & evidence	17%	25%	41%	23%
Fit, compliance & cost	40%	8%	21%	26%
Teaching & learning	18%	32%	3%	18%
Day-to-day use	18%	32%	21%	25%
Vendor & roadmap	7%	4%	15%	7%

Share of post-its in each timeframe, by theme area (analyst labels).

Source: workshop post-its (189 unique questions), EdTech Impact Summit, October 2026

Figure 3. Question themes by timeframe

	Short	Medium	Long	Not assigned	Total
Evidence of impact & outcomes	9	4	9	3	25
Data, reporting & insight	5	7	5	2	19
Integration, devices & tech fit	15	2	1	·	18
Features & functionality	4	6	1	5	16
Implementation, training & support	6	4	1	2	13
Pedagogy & teaching practice	·	12	1	·	13
Purpose, positioning & roadmap	6	2	5	·	13
Time saving & workload	4	5	3	·	12
Curriculum, policy & context fit	5	1	3	3	12
Pupil experience & engagement	8	4	·	·	12
Inclusion, SEND & differentiation	7	1	·	1	9
Data protection, safety & wellbeing	7	1	1	·	9
Usage & dosage	1	4	2	·	7
AI role, accuracy & safeguards	3	·	2	1	6
Cost & value	4	·	·	1	5

Number of post-its per theme. The Total column is shaded on its own scale.

Source: workshop post-its (189 unique questions), EdTech Impact Summit, October 2026

Figure 4. Each workshop drew a different mix of questions

	Tes Provision Map (n=25)	LbQ (n=83)	Magma Maths (n=43)	Zero Gravity (n=38)	All (n=189)
Evidence of impact & outcomes	8%	17%	14%	8%	13%
Data, reporting & insight	12%	5%	9%	21%	10%
Integration, devices & tech fit	24%	7%	12%	3%	10%
Features & functionality	4%	13%	7%	3%	8%
Implementation, training & support	8%	11%	5%	·	7%
Pedagogy & teaching practice	4%	5%	19%	·	7%
Purpose, positioning & roadmap	8%	8%	·	11%	7%
Time saving & workload	20%	5%	5%	3%	6%
Curriculum, policy & context fit	4%	8%	2%	8%	6%
Pupil experience & engagement	·	2%	7%	18%	6%
Inclusion, SEND & differentiation	·	5%	5%	8%	5%
Data protection, safety & wellbeing	·	4%	7%	8%	5%
Usage & dosage	4%	·	5%	11%	4%
AI role, accuracy & safeguards	4%	4%	5%	·	3%
Cost & value	·	6%	·	·	3%

Share of each workshop's post-its by theme (%). Shares are used because workshops produced different numbers of post-its.

Source: workshop post-its (189 unique questions), EdTech Impact Summit, October 2026

Figure 5. The five theme areas, by workshop

	Tes Provision Map	LbQ	Magma Maths	Zero Gravity
Impact & evidence	32%	22%	23%	21%
Fit, compliance & cost	32%	29%	26%	18%
Teaching & learning	4%	12%	30%	26%
Day-to-day use	24%	29%	21%	24%
Vendor & roadmap	8%	8%	·	11%

Share of each workshop's post-its (%). Area names are analyst labels.

Source: workshop post-its (189 unique questions), EdTech Impact Summit, October 2026

Workshop	Post-its	Top theme	Distinctive focus
Tes Provision Map	25	Integration, devices & tech fit (24%, 6 of 25)	Integration with existing systems and proof of measured time saving
LbQ (Learning by Questions)	83	Evidence of impact & outcomes (17%, 14 of 83)	Hard evidence of impact, plus practical features and set-up
Magma Maths	43	Pedagogy & teaching practice (19%, 8 of 43)	Pedagogy and unintended consequences in the classroom
Zero Gravity	38	Data, reporting & insight (21%, 8 of 38)	Teacher data use, pupil experience and dosage

The *top three themes* in each timeframe

For each timeframe, the three most common themes are shown below with an example post-it. The remaining themes are in Figure 3 and listed in full, with examples, in Appendix A2.

Short term: 84 post-its (44%)

Short-term questions were the most numerous and the broadest, covering 14 of the 15 themes. Although the prompt asked about potential impacts, the largest single theme was integration and devices, and questions about fit, compliance and cost made up 40% (34 of 84). Evidence of impact was already the second theme, and cost was raised on post-its only in the LbQ workshop.

15
post-its · 18%
8 groups

Integration, devices & tech fit

How the product connects to MIS, SSO, local authority data and other systems, which devices it runs on, and whether pupils need 1:1 devices.

Tes, LbQ, Magma, Zero Gravity

> *Will IT teams be happy? SSO, MIS integration, DPIA/GDPR compliant*
>
> LbQ Group 2 · post-it #8

9
post-its · 11%
6 groups

Evidence of impact & outcomes

Whether evidence already exists that the product works, for whom and through what mechanism, including requests for RCTs and for schools using it well.

Tes, LbQ, Magma, Zero Gravity

> *What is mechanism by which it impacts pupil outcomes? And what evidence supports that?*
>
> LbQ Group 4 · post-it #1

8
post-its · 10%
6 groups

Pupil experience & engagement

What pupils see and do, and what happens when they finish early, get stuck or go off task.

LbQ, Magma, Zero Gravity

> *What happens if student keeps saying "I don't know" i.e. task is pitched too high*
>
> Zero G Unknown Group · post-it #9

11 further themes in this timeframe: see Figure 3 and Appendix A2.

Medium term: 53 post-its (28%)

Medium-term questions (53, covering 13 themes) moved into the classroom. Pedagogy and teaching practice was the top theme (23%, 12 of 53), and 11 of those 12 came from Magma groups and LbQ Group 4. Together, day-to-day use and teaching and learning made up 64% (34 of 53). This is also where most usage and dosage questions sat (4 of 7).

12
post-its · 23%
5 groups

Pedagogy & teaching practice

How the product changes teaching: unintended consequences, deskilling, productive struggle, where it fits in a lesson and which pedagogies emerge.

Tes, LbQ, Magma

> *What are the unintended consequences of this product in the hands of a weak teacher?*
>
> Magma Group 2 · post-it #5

7
post-its · 13%
4 groups

Data, reporting & insight

Report options and analytics, parent access, learning maps, class and trust-level insight, and whether teachers act on the data.

Tes, LbQ, Magma, Zero Gravity

> *How will you ensure teachers will use the data and action it?*
>
> Zero G Unknown Group · post-it #11

6
post-its · 11%
4 groups

Features & functionality

Home learning, revision, self-assessment and open-ended questions.

Tes, LbQ, Magma

> *Can pupils self-assess?*
>
> LbQ Group 2 · post-it #10

10 further themes in this timeframe: see Figure 3 and Appendix A2.

Long term: 34 post-its (18%)

Long-term questions were the fewest (34, covering 12 themes). Evidence of impact led (26%, 9 of 34), followed by roadmap and data use (5 each). Several long-term evidence questions concern effects a 3 to 6 month testbed cannot answer, such as long-term learning and KS2 results, which fits the brief's steer away from long-term questions for the worksheet.

9
post-its · 26%
5 groups

Evidence of impact & outcomes

Long-term learning, effects on KS1, MTC and KS2 results, evidence for mathematical thinking, and whether the vendor has a theory of change.

Tes, LbQ, Magma

> *What is the effect on long-term learning?*
> **LbQ Group 4 · post-it #17**

5
post-its · 15%
3 groups

Data, reporting & insight

Seeing progress over time, school-level overviews and using data outside the platform.

LbQ, Magma, Zero Gravity

> *Teacher dashboard: Can a teacher easily see progress over time?*
> **Zero G Group 1 · post-it #11**

5
post-its · 15%
4 groups

Purpose, positioning & roadmap

Development roadmap, scaling strategy, support for more subjects, alternatives and resilience to another lockdown.

Tes, LbQ, Zero Gravity

> *What is your dev. roadmap?*
> **LbQ Group 1 · post-it #18**

9 further themes in this timeframe: see Figure 3 and Appendix A2.

Not assigned: 18 post-its (10%)

Eighteen post-its (10%) were not sorted into a timeframe, all from LbQ groups (11 from Group 4). They lean towards product mechanics and curriculum fit.

5
post-its · 28%
3 groups

Features & functionality

Question mechanics: re-testing, anti-cheating, question versions, home use and tailoring.

LbQ

> *Can students click off the Qu to use other systems to find the answers? Anticheating?*
> **LbQ Group 1 · post-it #19**

3
post-its · 17%
2 groups

Curriculum, policy & context fit

Fit with different curricula and subjects not yet covered.

LbQ

> *Can it fit different curricula?*
> **LbQ Group 4 · post-it #23**

3
post-its · 17%
1 group

Evidence of impact & outcomes

Evidence that learning is happening, and the extent of improvement.

LbQ

> *Evidence that learning is happening?*
> **LbQ Group 4 · post-it #25**

5 further themes in this timeframe: see Figure 3 and Appendix A2.

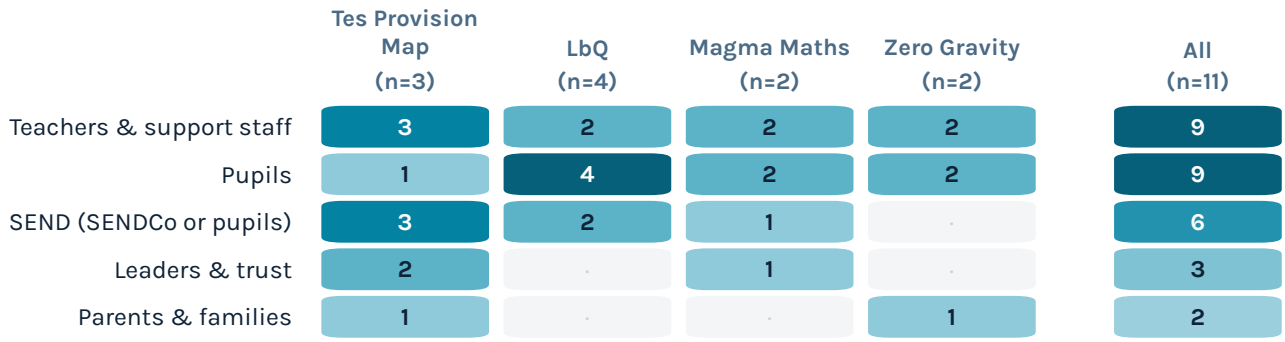
How groups would design *an evaluation*

Asked to sketch a 3 to 6 month evaluation in about 20 minutes, most groups proposed around six months, attainment plus teacher and pupil voice, and a pre/post baseline or a comparison group. No group wanted the product company to draw the conclusions from the data collected and analysed, and most wanted pupils' and parents' voices to be heard.

Each section opens with a heatmap showing how many groups in each workshop chose each option, followed by what was most common and where groups differed. Columns show the number of groups that answered (n), and the final column gives the total across all workshops. Free-text fields were coded into categories, which can be checked in the "Worksheet coding" tab of the coded workbook. Sections left empty are excluded from the counts. Several worksheets were only partly completed (Magma Groups 1 to 3 and Zero Gravity Group 1), which may reflect the 20 minutes available, but may also show that school teams need support to design an evaluation (see recommendation 4).

Context

Figure 6. Types of users the product supports



Answered on 11 of 13 worksheets (left blank by Magma Groups 1 and 2).

Source: workshop worksheets (13 worksheets from 14 groups), EdTech Impact Summit, October 2026

MOST COMMON

Teachers or other staff (9 of 11) and pupils (9 of 11). SEND was named on 6 of 11, either as SENDCo users or as pupils with SEND. Parents or families were named twice (Tes Group 1, Zero Gravity Group 1) and leaders or trust teams three times (Tes Groups 1 and 2, Magma Group 3).

WHERE GROUPS DIFFERED

The Tes groups described a staff tool (SENDCo, DSL, teachers, leaders), while the maths products were described through pupils and year groups.

Figure 7. Users for this evaluation

	Tes Provision Map (n=3)	LbQ (n=3)	Magma Maths (n=2)	Zero Gravity (n=1)	All (n=9)
Single class	1	·	·	·	1
One or two year groups	·	2	·	1	3
Pupils across several year groups	·	1	1	·	2
Several schools	1	·	·	·	1
SEND focus	·	1	1	·	2
Named staff roles	2	1	1	·	4
Numbers given	1	2	1	·	4

Answered on 9 of 13 (left blank by LbQ Group 4, Zero Gravity Group 1 and Magma Groups 1 and 2).

Source: workshop worksheets (13 worksheets from 14 groups), EdTech Impact Summit, October 2026

MOST COMMON

The worksheet asked groups to be specific about numbers, and 4 of 9 did: LbQ Group 1 (a year group of 2 to 3 classes of 30, with 25% SEND and 5% EHCP), Tes Group 2 (one SENDCo, one leader and one class), Magma Group 4 (120 pupils, 6 teachers, 2 subject leads and an academic head) and LbQ Group 3 (Years 3 and 4 in a two-form-entry school).

WHERE GROUPS DIFFERED

Scale ranged from one class to "multiple pilot schools" (Tes Group 1). Other answers named roles only (for example "SLT & SENDCO") or a sample (Magma Group 3: "sample of children with SEND in each year group").

Figure 8. If pupils are included, what ages

	Tes Provision Map (n=3)	LbQ (n=4)	Magma Maths (n=1)	Zero Gravity (n=2)	All (n=10)
Primary (Years 3 to 6)	·	4	1	·	5
Secondary (KS3 to KS4)	1	·	·	2	3
All ages	2	·	·	·	2
Not applicable (staff tool)	1	·	·	·	1

Answered on 10 of 13 (left blank by Magma Groups 1 to 3).

Source: workshop worksheets (13 worksheets from 14 groups), EdTech Impact Summit, October 2026

MOST COMMON

Primary, Years 3 to 6 (5 of 10: all four LbQ groups and Magma Group 4). Three were secondary (Tes Group 3 Years 9 and 10, Zero Gravity KS3 and Year 9). Tes Group 1 wrote "All" and Tes Group 2 "N/A", likely reflecting a staff-facing tool.

WHERE GROUPS DIFFERED

Age appeared to follow the product, with the two maths products in primary (Magma's research questions also named upper KS2 and Year 5) and Zero Gravity in KS3.

Evidence base

Figure 9. Data points required for the evaluation

	Tes Provision Map (n=3)	LbQ (n=4)	Magma Maths (n=3)	Zero Gravity (n=2)	All (n=12)
Attainment & progress assessments	2	4	2	2	10
Teacher & staff voice	2	4	2	1	9
Baseline / pre-post measure	2	3	2	1	8
Pupil voice	·	4	3	1	8
Usage & adoption data	·	2	2	·	4
Existing school data (MIS, prior cohorts)	2	·	1	·	3
Observation & implementation quality	1	2	·	·	3
Engagement, behaviour & wellbeing	1	2	·	·	3
Time & workload	2	·	·	·	2
Leader voice	1	·	1	·	2
Subgroup analysis (SEND, FSM)	·	2	·	·	2
Parent voice	·	1	1	·	2
Compliance & security checks	1	·	·	·	1
Trial design (e.g. micro-RCT)	·	1	·	·	1
Knowledge retention	·	1	·	·	1

Answered on 12 of 13 (left blank by Magma Group 2).

Source: workshop worksheets (13 worksheets from 14 groups), EdTech Impact Summit, October 2026

MOST COMMON

Attainment or progress assessments (10 of 12), teacher and staff voice (9), a baseline or pre/post measure (8) and pupil voice (8). Lists held about five items on average, from 1 (Zero Gravity Group 1) to 7 (Tes Group 2, LbQ Groups 2 and 3).

WHERE GROUPS DIFFERED

Usage data (4 of 12), existing school data such as MIS or prior cohorts (3), observation or learning walks (3) and engagement, behaviour or wellbeing (3) were less common. Single mentions included a time and workload baseline (Tes), compliance checks (Tes Group 2), micro-RCTs (LbQ Group 3) and knowledge retention (LbQ Group 4). The worksheet asked "what and when", but only about half the lists said when, usually as baseline and end-point.

Figure 10. Who collects the data

	Tes Provision Map (n=2)	LbQ (n=4)	Magma Maths (n=2)	Zero Gravity (n=1)	All (n=9)
Internal team	2	4	1	·	7
Product company	1	3	·	1	5
Researcher	·	·	1	1	2
Independent org	1	1	1	·	3

Answered by 9 groups.

Source: workshop worksheets (13 worksheets from 14 groups), EdTech Impact Summit, October 2026

MOST COMMON

An internal team (7 of 9), then the product company (5 of 9). An independent organisation was ticked by 3 groups and a researcher by 2. Eight of the 9 groups ticked two options, suggesting most saw collection as shared.

WHERE GROUPS DIFFERED

All four LbQ groups chose an internal team, and three of them also the product company. Tes Group 1 crossed out the product company and chose an independent organisation instead, and Magma Group 4 wrote that the school and the product company should agree on an independent organisation.

Figure 11. Who analyses the data

	Tes Provision Map (n=1)	LbQ (n=4)	Magma Maths (n=2)	Zero Gravity (n=1)	All (n=8)
Internal team	1	3	·	1	5
Product company	·	2	·	·	2
Researcher	·	1	1	1	3
Independent org	·	1	1	·	2
MAT / trust team	·	1	·	·	1

Answered by 8 groups.

Source: workshop worksheets (13 worksheets from 14 groups), EdTech Impact Summit, October 2026

MOST COMMON

An internal team (5 of 8), then a researcher (3), an independent organisation (2) and the product company (2, both LbQ). One group added a MAT team.

WHERE GROUPS DIFFERED

LbQ Group 1 spread analysis across four parties (internal team, researcher, product company and MAT team), while LbQ Group 3 and Magma Group 4 gave it to an independent organisation alone and Magma Group 2 to a researcher alone. LbQ Group 4 made small marks in the researcher and independent organisation boxes, which are not counted as ticks.

Figure 12. Who formulates the conclusions

	Tes Provision Map (n=1)	LbQ (n=4)	Magma Maths (n=2)	Zero Gravity (n=0)	All (n=7)
Internal team	1	3	·	·	4
School leadership	1	3	·	·	4
MAT / trust team	·	1	·	·	1
Researcher	·	·	2	·	2
Independent org	·	1	·	·	1
Product company	·	·	·	·	·

Answered by 7 groups; no Zero Gravity group completed this field.

Source: workshop worksheets (13 worksheets from 14 groups), EdTech Impact Summit, October 2026

MOST COMMON

An internal team (4 of 7) and school leadership (4 of 7), an option offered only at this stage. No group wanted the product company to draw the conclusions from the data collected and analysed. In total, 71% (5 of 7) gave conclusions to the school or trust, LbQ Group 3 alongside an independent organisation.

WHERE GROUPS DIFFERED

Both Magma groups that answered chose a researcher only. LbQ Group 3 paired its internal team with an independent organisation, and LbQ Group 4 crossed out the product company.

Figure 13. Evaluation team: who needs to give input

	Tes Provision Map (n=3)	LbQ (n=4)	Magma Maths (n=2)	Zero Gravity (n=0)	All (n=9)
Headteacher or senior leader	3	4	1	·	8
Subject lead	1	4	2	·	7
Class teachers	3	3	1	·	7
SENCO	3	3	·	·	6
Data protection lead (DPO)	2	4	1	·	7
IT or network lead	2	4	1	·	7
Safeguarding lead	1	2	·	·	3
Budget holder	2	2	1	·	5

Answered by 9 groups.

Source: workshop worksheets (13 worksheets from 14 groups), EdTech Impact Summit, October 2026

MOST COMMON

Headteacher or senior leader (8 of 9), then class teachers, subject lead, DPO and IT lead (7 each), SENCO (6) and budget holder (5). Two groups (LbQ Groups 1 and 2) ticked all eight roles.

WHERE GROUPS DIFFERED

Safeguarding lead was the least ticked (3 of 9). LbQ Group 4 explicitly crossed out class teachers, safeguarding lead and budget holder. Magma Group 2 ticked only subject lead and class teachers. Four worksheets wrote pupils, parents or the MAT into "Other". No Zero Gravity group completed this field.

Figure 14. Whose voices and experiences should be included

	Tes Provision Map (n=3)	LbQ (n=4)	Magma Maths (n=2)	Zero Gravity (n=1)	All (n=10)
Pupils	2	4	2	1	9
Teachers	3	4	2	1	10
Parents	3	2	2	·	7
Support staff	3	1	·	·	4
School leadership	2	1	1	1	5

Answered by 10 groups.

Source: workshop worksheets (13 worksheets from 14 groups), EdTech Impact Summit, October 2026

MOST COMMON

Teachers (10 of 10) and pupils (9 of 10), then parents (7), school leadership (5) and support staff (4).

WHERE GROUPS DIFFERED

All three Tes groups included support staff and parents, while LbQ Groups 3 and 4 kept to pupils and teachers. Write-ins added the SENDCo, trust leaders and governing board, and the MAT.

Figure 15. How should findings be presented

	Tes Provision Map (n=3)	LbQ (n=3)	Magma Maths (n=2)	Zero Gravity (n=1)	All (n=9)
Dashboard	1	3	1	1	6
Written report	·	2	2	1	5
One-pager	2	1	·	·	3
Case study	2	·	1	·	3
Live session	·	·	·	1	1

Answered by 9 groups.

Source: workshop worksheets (13 worksheets from 14 groups), EdTech Impact Summit, October 2026

MOST COMMON

Dashboard (6 of 9) and written report (5), then one-pager (3), case study (3) and live session (1). Six of 9 groups chose two or more formats.

WHERE GROUPS DIFFERED

The Tes groups preferred light-touch formats (one-pagers and case studies), while LbQ and Zero Gravity favoured dashboards. LbQ Group 1 noted a one-pager "for LGB + exec", showing outputs were seen as audience-specific.

Wrapping up: context and evidence base

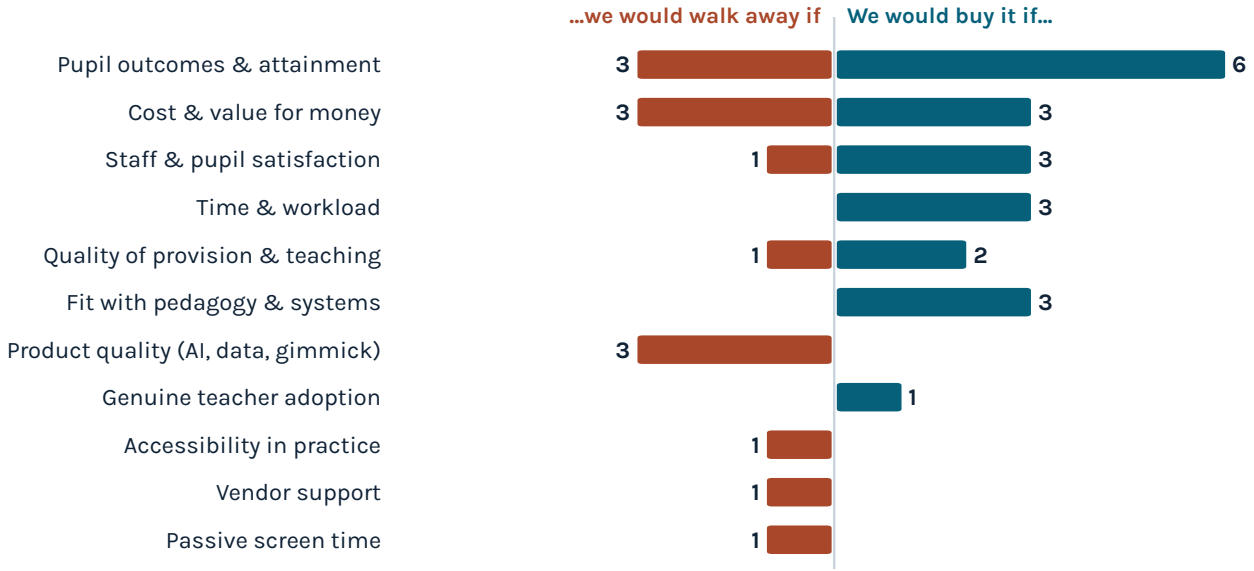
Taken together, the Context and Evidence base sections give a consistent picture. Groups pictured evaluations built around teachers and pupils, mostly in primary maths, with SEND named on over half of worksheets (6 of 11). Attainment, teacher voice, a baseline and pupil voice formed the core evidence, while usage data was listed by only a third of groups (4 of 12).

Groups were open to the product company supplying data (5 of 9) but no group wanted it to draw conclusions from the data collected and analysed (0 of 7). Most groups gave conclusions to the school or trust (5 of 7). Senior leaders, the DPO and the IT lead were the roles most often ticked for the evaluation team, and teachers, pupils and parents as the voices to hear. Groups most often wanted findings as dashboards and written reports, often in more than one format for different audiences.

Much of the variation lined up with the product (though, with one product per workshop, this cannot be separated from who was in each room): staff-facing evaluations for Tes, primary maths for LbQ and Magma, KS3 for Zero Gravity, and a stronger preference for external researchers in the Magma groups. The next two sections cover how groups would decide whether to buy, and the research questions they drafted.

Decision

Figure 16. What would make groups buy the product or walk away



Number of statements in the worksheet decision section, coded by theme. Completed on 7 of 13 worksheets.

Source: workshop worksheets (13 worksheets from 14 groups), EdTech Impact Summit, October 2026

We would buy it if... (positive impact)

Completed on 7 of 13 worksheets (21 statements).

MOST COMMON

Improved pupil outcomes (6 of 21 statements, from 5 groups). Then value for money or affordability, time saving, staff and pupil satisfaction, and fit with pedagogy or systems (3 statements each).

WHERE GROUPS DIFFERED

Tes groups framed success around provision quality, affordability and time; LbQ and Magma around attainment and progress. Only two statements hinted at a threshold ("positive impact that is quantified", "noticeable academic growth"), and none set a figure.

We would walk away if... (negative or no impact)

Completed on 7 of 13 worksheets (14 statements).

MOST COMMON

No improvement in outcomes, cost or value concerns, and product quality (3 statements each). Product quality covered "partial data output", "AI generic outcome" and the product turning out to be "a gimmick".

WHERE GROUPS DIFFERED

Single mentions covered provision not improving, accessibility not working in practice, lack of vendor support, negative user feedback, and passive screen time. Magma Group 4 completed the positive side only, and both Zero Gravity worksheets left the decision section blank.

Research question

Product, cohort and setting

All 13 named the product; 13 gave a cohort and 12 a setting.

MOST COMMON

A primary year group or key stage (7 of 13: all four LbQ groups and Magma Groups 1, 2 and 4). Two were secondary (Zero Gravity), one focused on pupils with SEND across year groups (Magma Group 3) and three on staff or schools (Tes).

WHERE GROUPS DIFFERED

Settings ranged from very specific ("a high deprivation/high EAL school", "our specialist resource base", "independent school") to generic ("primary school", "1 school"). Zero Gravity Group 1 left the setting blank, and the unknown Zero Gravity group wrote "Business".

Duration

Answered on 12 of 13.

MOST COMMON

About six months or two terms (7 of 13), at the top of the 3 to 6 month window set by the brief. Two gave a range within the window (3 to 6 and 4 to 6 months).

WHERE GROUPS DIFFERED

Three went beyond the window to a year (Tes Groups 1 and 2, Magma Group 3), and Zero Gravity Group 1 left it blank.

Specific outcome

Answered on 13 of 13.

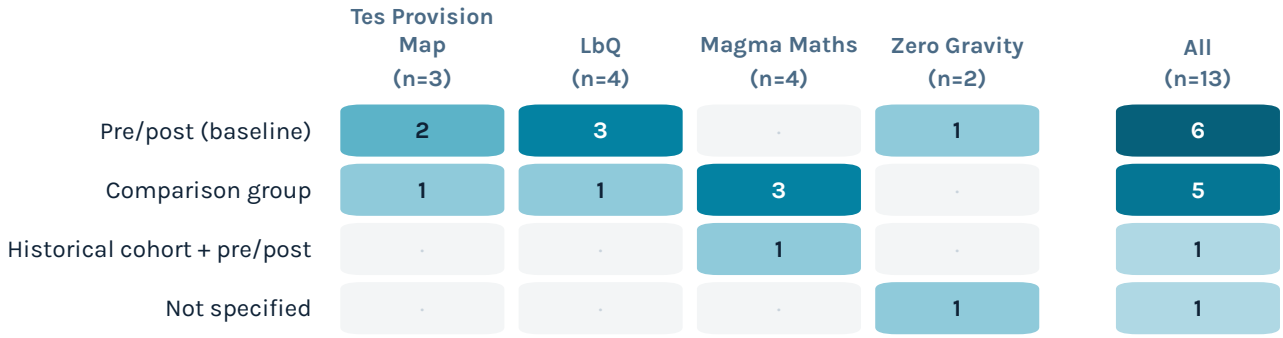
MOST COMMON

Attainment or academic progress (5 of 13), with two more on knowledge retention or knowledge and confidence (LbQ Group 4, Magma Group 3).

WHERE GROUPS DIFFERED

Two Magma groups chose resilience in maths, LbQ Group 2 combined outcomes and engagement, and the Tes groups chose workload, reporting time and quality of provision, or successful implementation in lessons.

Figure 17. Data source and comparison



Data source answered on 11 of 13; comparison on 12 of 13.

Source: workshop worksheets (13 worksheets from 14 groups), EdTech Impact Summit, October 2026

MOST COMMON

Assessment or test scores as the data source (7 of 11) and a pre/post baseline as the comparison (6 of 13).

WHERE GROUPS DIFFERED

5 of 13 proposed a comparison group (a control class, business as usual, or non-users) and Magma Group 1 a historical cohort. Less common data sources were time and workload measures (Tes Group 1) and counts of teacher interventions and pupil self-corrections (Magma Group 2).

All research questions

Workshop	Group	Cohort	Duration	Outcome	Comparison
Tes Provision Map	Group 1	Pilot Schools	1 Year	Education outcomes	Pre/post (baseline): Current state
Tes Provision Map	Group 2	1 class / Sendco	1 year or 2 terms	reporting time + Q of provision	Pre/post (baseline): the baseline
Tes Provision Map	Group 3	Senco	2 terms	Successful implementation in lesson	Comparison group: Pre / Post test scores.
LbQ	Group 1	Year 4 students (focus on SEND)	3-6 months	student data in maths	Pre/post (baseline): baseline.
LbQ	Group 2	Pupils in KS2	6 Months	outcomes and engagement	Pre/post (baseline): baseline outcomes
LbQ	Group 3	Year 3 / Year 4	4-6 months	attainment defined by pre, post test	Comparison group: business as usual.
LbQ	Group 4	Key Stage 2	6 months	knowledge retention & application	Pre/post (baseline): baseline.
Magma Maths	Group 1	UKS2	two full terms	student resiliance	Historical cohort + pre/post: prior cohorts + before/after attitudes
Magma Maths	Group 2	a Year 5 class	six months	student resilience in Maths	Comparison group: a Year 5 class in the same school (with ideally the same teacher) that didn't use the tool
Magma Maths	Group 3	children with SEND (specific)	a year	on pupil knowledge + confidence	Comparison group: peers who have not used Magma Math
Magma Maths	Group 4	7-9 years	6 months	academic growth	Comparison group: non users
Zero Gravity	Group 1	KS3	Not stated	attainment in maths + history	Not specified
Zero Gravity	Unknown Group	Y9	6 months	attainment	Pre/post (baseline): baseline assess

Where vendor claims met *participants' questions*

Each workshop opened with a 10-minute vendor presentation and live demo, with vendors briefed to end on about three claims about impact. The claims below are taken from the decks LbQ and Zero Gravity provided before the workshop as the content for those 10 minutes. We are unable to confirm whether every claim was presented in each workshop, so the tables show how participants' questions line up with each vendor's stated claims, rather than how groups reacted to them.

The evidence data each vendor already holds was presented later, after the worksheets were completed, by design (see page 2), so groups planned their evaluations without knowing what that data was. LbQ was a partial exception: it covered some of its evidence at the start and added more in the evidence slot. Presentation decks were requested from all four vendors before the workshop; Tes and Magma Maths have been asked for a short summary, which will be added when received.

Zero Gravity

WHAT THE VENDOR PRESENTED (COMPANY-SUPPLIED)

An AI tutor for secondary schools (Key Stages 3 to 5, Years 7 to 13), built on Oak National Academy at KS3 and exam board specifications at KS4 and KS5. Teachers set one assignment a week, the tutor coaches each pupil through it, the teacher reviews misconceptions and the next assignment targets the gaps.

EVIDENCE SHOWN

An "evidence ladder" from the vendor's own data to independent research: usage data (live in every school), session quality, pupil and teacher voice, results by usage (first cohort, August 2026: consistent users +0.5 grades, light users +0.1), classroom observation, and a randomised trial in development with a UK university. 56% of active pupils return week on week. The deck also sets out its foundations: data hosted in the UK, pupil data never used to train a model, built to the DfE generative AI product safety standards, and safeguarding by design.

CLAIM **Adaptability: one teacher action creates 30 different learning experiences; by the next lesson the teacher has the class's top three misconceptions, ranked by pupil count, and re-teaches at least one.**

WHAT GROUPS ASKED

> *What do you mean by student misconceptions? How do you see them?*
>
> Zero G Group 1 · post-it #12

> *How do you know the teachers have enough data literacy to interpret the figures they are seeing?*
>
> Zero G Group 1 · post-it #3

> *How will you ensure teachers will use the data and action it?*
>
> Zero G Unknown Group · post-it #11

WHAT GROUPS PLANNED TO MEASURE

No worksheet listed a measure of what teachers do with the misconception data.

OVERLAP AND GAPS

The deck shows the misconception ranking is live. Whether teachers act on it, the step the claim depends on and the one groups questioned, is not yet measured by vendor or groups.

CLAIM **Continuity: in a normal lesson, cover, isolation or an unfilled post, pupils complete as much and score as well as in a taught lesson, and every hour lands in the markbook.**

WHAT GROUPS ASKED

> *Interopeability with MIS? In terms of markbooks?*
>
> Zero G Unknown Group · post-it #6

> *Is this best used in class or @ home?*
>
> Zero G No-Worksheet Group · post-it #1

WHAT GROUPS PLANNED TO MEASURE

Not addressed.

OVERLAP AND GAPS

Few questions related to this claim, which may not have been covered in detail on the day. Testing it would need a comparison of cover and taught lessons.

CLAIM **Attainment: pupils who complete ten or more sessions between assessments move up at least half a grade more than those who complete fewer than three.**

WHAT GROUPS ASKED

> *Attain. What if someone is at the top grade? or bottom grade -> 0.5 not enough*
>
> Zero G Group 1 · post-it #9

> *Attainment dosage correlation? => context of the numbers*
>
> Zero G Group 1 · post-it #8

> *How often does a student need to use the tool to reach your goal*
>
> Zero G Unknown Group · post-it #14

WHAT GROUPS PLANNED TO MEASURE

Both worksheets chose attainment (KS3 maths and history; Year 9 mocks against a baseline), with pupil and teacher voice. Neither proposed a comparison group.

OVERLAP AND GAPS

Groups questioned whether half a grade is meaningful at the top and bottom of the range and asked for the context behind a dose-response figure. Comparing heavy and light users can reflect which pupils are more motivated; the planned trial could help address this.

CLAIM **Foundations: UK data hosting, no model training on pupil data, DfE generative AI product safety standards, safeguarding by design.**

WHAT GROUPS ASKED

> *How does the safeguarding function work?*
>
> Zero G Unknown Group · post-it #1

> *Where is data stored?*
>
> Zero G Unknown Group · post-it #2

> *Safety. GDPR. Teacher/Student training. Data collection to evaluate.*
>
> Zero G No-Worksheet Group · post-it #6

WHAT GROUPS PLANNED TO MEASURE

Not a worksheet focus.

OVERLAP AND GAPS

The deck appears to answer most of the safety and data questions groups raised in the short term.

Two post-its asked about primary use (Zero Gravity No-Worksheet Group #5 and #7), which is outside the product's stated Key Stage 3 to 5 range.

LbQ (Learning by Questions)

WHAT THE VENDOR PRESENTED (COMPANY-SUPPLIED)

The deck focuses on Springboard Diagnostics, which identifies gaps in pupils' prior knowledge at whole-class, group and individual level and provides resources to address each gap. LbQ described an evidence cycle of design, use and impact, with a theory of change grounded in EEF guidance on prior knowledge and misconceptions, and referenced the DfE and Chartered College "six signs of good evidence".

EVIDENCE SHOWN

Two hypotheses: improved academic outcomes and reduced teacher workload. Evidence so far: teacher surveys, case studies and an external review, with pupil response data used to inform product development.

CLAIM Improves academic outcomes, by identifying and closing gaps in prior knowledge.

WHAT GROUPS ASKED

> *What is mechanism by which it impacts pupil outcomes? And what evidence supports that?*
>
> LbQ Group 4 · post-it #1

> *Do LBQ have examples of RCT's/evidence to back up impact?*
>
> LbQ Group 3 · post-it #5

> *Which user/group has the tool/s been shown to work to improve outcomes most often?*
>
> LbQ Group 3 · post-it #3

WHAT GROUPS PLANNED TO MEASURE

All four worksheets chose attainment in KS2 maths or knowledge retention, with a baseline; one used a business-as-usual comparison and micro-RCTs, and one focused on pupils with SEND.

OVERLAP AND GAPS

Groups asked for, and designed, causal evidence beyond the surveys, case studies and review LbQ currently reports. Its theory of change explains how the product is meant to work, but does not yet show evidence that it does.

CLAIM Reduces teacher workload.

WHAT GROUPS ASKED

> *How do LBQ understand Teacher time saved?*
>
> LbQ Group 3 · post-it #7

> *How much effort is required to set questions as a teacher.*
>
> LbQ Group 1 · post-it #7

> *How do teachers reallocate their time?*
>
> LbQ Group 4 · post-it #10

WHAT GROUPS PLANNED TO MEASURE

Time saving appeared only as a reason to buy (LbQ Groups 1 and 4); no worksheet planned to measure teacher time.

OVERLAP AND GAPS

The claim is evidenced by teacher surveys, and groups did not plan a time measure either. A timed baseline would test it directly.

Groups discussed LbQ as a whole platform, while the deck focuses on one feature, Springboard Diagnostics.

What this may suggest

The claims were specific and testable, and groups' questions and plans lined up most with the attainment claims. The evidence offered sits mainly on the lower rungs (usage data, surveys and case studies; Zero Gravity has a randomised trial in development and LbQ cites an external review), while groups' own plans asked for attainment against a baseline and, in some LbQ groups, a comparison class. Claims about what teachers do (re-teaching from misconception data, saving time) were raised on post-its but built into no evaluation plan, and few questions related to Zero Gravity's continuity claim. Both point to mapping each claim to a measure before a testbed starts.

Further testbed design learnings *informed by the workshops*

A shared core plus product-specific modules

Beyond the ten learnings, the workshops point to one design principle for testbeds. Every evaluation could share a common core: the same checks, measures and agreements for every product, so that results can be compared across products and schools know what to expect. Alongside the core sits a smaller product-specific module, chosen for the type of product and the claims it makes.

The workshop data points both ways. Six of the 15 themes came up in all four workshops (evidence of impact, integration, data and reporting, features, curriculum fit and time saving), and together they account for 54% (102 of 189) of post-its. On the worksheets, attainment, teacher voice and a baseline appeared in data-point lists for all four vendors, and pupil voice for three. At the same time, each workshop had a distinctive focus that ran at two to four times the overall share: integration (24% against 10% overall) and time saving (20% against 6%) for Tes, pedagogy for Magma (19% against 7%), and data and reporting (21% against 10%), pupil experience (18% against 6%) and usage (11% against 4%) for Zero Gravity. LbQ's profile was the broadest and was the only workshop where cost was raised on post-its (5). The research questions diverged in the same way: workload and quality of provision for Tes, resilience and comparison classes for Magma, KS3 attainment for Zero Gravity, and KS2 maths with SEND subgroups for LbQ.

What the core could contain

1 **Fit and compliance check before the pilot starts**
Answers: Integration, devices, DPIA, accessibility and safeguarding
Integration was the top short-term theme (18%, 15 of 84) in three of the four workshops (all except Zero Gravity); 7 of 9 groups put the DPO and IT lead on the evaluation team

2 **Vendor evidence pack and theory of change**
Answers: The claim, the mechanism, intended dose and who it is for
Evidence of impact was the top theme overall (13%, 25 of 189) and came up in all four workshops

3 **Fidelity and process measures**
Answers: Is the product used as intended, and do teachers act on what it shows?
Groups raised these questions on post-its (for example, whether teachers act on misconception data) but rarely planned to measure them; in a three to six month window they may show change sooner than attainment

4 **Teacher and pupil voice**
Answers: How was it experienced?
Teacher voice in 9 of 12 lists and pupil voice in 8 of 12; pupils ticked as a voice by 9 of 10 groups

5 **Usage data**
Answers: Was it used as intended, and how much?
Only 4 of 12 lists, but usage and dosage questions came up in three of the four workshops

6 **Workload measure**
Answers: Did it save or cost teacher time?
Time saving was raised in all four workshops (12 post-its)

7 **Agreed roles and decision criteria**
Answers: Who analyses, who concludes, and what would lead to buying or walking away
No group wanted the product company to draw conclusions from the data collected and analysed (0 of 7); buy and walk-away criteria were set on 7 of 13 worksheets

What modules could look like

The suggested module measures are our proposals based on what groups asked, not something groups wrote.

Tes (Provision Map)	**Groups focused on:** Integration with MIS and LA data; how time savings were measured **Possible module:** Timed baseline of SENDCo and teacher reporting tasks; data-flow check; quality-of-provision review
LbQ	**Groups focused on:** Evidence and mechanism; features and set-up; cost **Possible module:** SEND and FSM subgroup analysis; implementation checks through learning walks; cost-per-pupil and value analysis
Magma Maths	**Groups focused on:** Pedagogy and unintended consequences in the classroom **Possible module:** Classroom observation and teacher interviews; comparison class; resilience or self-correction measure
Zero Gravity	**Groups focused on:** Teacher data use, pupil experience and dosage **Possible module:** Dosage-to-outcome analysis; tracking whether teachers view and act on data; pupil survey on progress awareness

Baseline and end-point outcome. Groups listed attainment in 10 of 12 data-point lists and a baseline in 8 of 12, for all four vendors. In a short (3 to 6 month) testbed this is better treated as part of the product-specific module, because the primary outcome differs by product and attainment may be slow to show change in that time. In a longer testbed it could become part of the core.

Modules are better defined by product type and claim than by vendor, so that two products of the same type are tested in the same way and the core stays comparable across all of them.

This rests on four products and small groups, so it is a working hypothesis to test in the next testbed cycle. One open question is how far the core can stay identical when primary outcomes differ, for example workload for an administrative tool and attainment for a maths product. One option is to fix the process measures (fit, usage, voice, roles and decision criteria) and let the primary outcome vary by product.

Designing testbeds schools *will trust*

1 Stage the evidence
Match evidence to when each question can be answered, not to its importance: a short-term fit pack (integration, DPIA, accessibility, cost), a medium-term implementation and pedagogy review, and a longer-term impact study.

2 Standardise a core protocol built on early signals
In a three to six month window, lead with fidelity and process measures (is the product used as intended, and do teachers act on its data?), plus usage data and teacher and pupil voice. Treat attainment as a product-specific measure where the timescale allows, not a requirement for every product.

3 Add a product-specific module
For example, timed workload baselines for admin tools, classroom observation for pedagogy-heavy tools, and dosage analysis for practice platforms.

4 Build evaluation support into the testbed
Researchers were in at least the LbQ, Magma Maths and Zero Gravity workshops, but there was no set way for them to take part, so some stepped away from the groups and some joined in. Groups moved quickly to assessment data as the measure, and several worksheets were only partly filled in. The 20 minutes available is likely to have played a part, but together these suggest school teams would benefit from help designing a testbed study. Practical evaluation support should therefore be a recommended part of what a testbed offers.

5 **Ask vendors for a theory of change up front**
Mechanism, intended dose, target users and known limitations, so schools can test claims rather than features. LbQ already presents one, which suggests it is workable.

6 **Keep interpretation independent**
Vendors can supply data, and 2 of 8 groups also gave them a share of the analysis, but none gave them the conclusions. Set out who collects, analyses and concludes in the testbed agreement.

7 **Define keep or drop criteria before the pilot**
Outcomes, value for money and product quality were the most mentioned factors. Agree thresholds at the start so decisions are not post hoc.

8 **Build teacher data literacy into implementation**
Two Zero Gravity groups questioned whether teachers would act on the data without support.

9 **Prepare for AI questions**
Six post-its across three vendors asked about AI accuracy, training data and pupils' understanding of AI. Vendors should have clear answers.

Learnings for the next workshops

Running four workshops in one day also taught us about the workshop format itself. These points are for the team running future co-design workshops.

1 **Brief the supporting person as fully as the facilitator**
Every workshop had a supporting person, but they were not the workshop designers. Worksheets were more complete where the designers were in the room, though group make-up may also play a part: LbQ (96%, 50 of 52 sections) and Tes (85%, 33 of 39), against Magma Maths (50%, 26 of 52) and Zero Gravity (46%, 12 of 26, with a third group producing no worksheet). The supporting person needs the same knowledge of the plan, scope and goal as the facilitator, so they can guide groups through every section rather than only hand out post-its and pens.

2 **Lead the closing discussion, not just the read-out**
The close was meant to compare vendor data with what schools wanted to measure, but in practice groups read out their research questions. A short structure would help: each group names the claim its question tests and the data it would need, the vendor says what it can already provide, and the facilitator draws out where the two match and where they differ. A note-taker should record the points of match and mismatch.

3 **Collect every vendor's slides before the day**
Slides were requested from all four vendors, but only LbQ and Zero Gravity provided them, so questions cannot yet be linked to the Tes and Magma Maths claims. Make the claims slide a condition of taking part.

4 **Set up a process for vendor staff in the room**
Vendors agreed beforehand to keep their staff out of the groups, but in the workshops some vendor staff joined in and were reluctant to step back, despite repeated nudges from the organising team. Agree a clear process in advance: vendor staff answer factual questions only, and if they join a group, the organiser has the right to ask them to leave the room.

5 **Decide how researchers in the room take part**
Researchers were in at least the LbQ, Magma Maths and Zero Gravity workshops, but with no set way to take part, some stepped back and some joined in. Decide in advance whether they observe or actively scaffold study design, and brief them accordingly.

6 **Capture who is in each group first, and record the workshops**
Groups were asked for names and roles as the last step, after the worksheet, and only some gave them, so differences between groups are hard to explain. Make this step 1 once groups are formed (a sign-in sheet per table) and update the session plan. With participants' consent, also record the workshops.

7 **Label materials by workshop and group before the day**
One Zero Gravity group number is unknown, and three post-its filed under Zero Gravity match post-its from other workshops word for word. Pre-numbered worksheets and post-it pads per table would make the data easier to trace.

APPENDIX · EVIDENCE AND METHOD

Data, methods and *full results*

This appendix sets out where each learning in the report comes from and how it was derived. A1 traces every figure to the transcription and coded workbooks; A2 lists every theme in each timeframe with an example post-it; A3 gives the method, sources and limitations; A4 lists the post-its excluded from the counts and how complete the worksheets were; A5 defines the themes; A6 shows the worksheet groups filled in.

A1 Finding trail: where each learning comes from

A2 Every theme by timeframe, with example post-its

A3 Method, sources and limitations

A4 Data notes: excluded post-its and worksheet completion

A5 Theme glossary

A6 The testbed worksheet

A1 · Finding trail: where each learning comes from

Two sources are referenced: the transcription workbook (Summit_workshops_transcription_Oct2026.xlsx, one tab per vendor, one block per group) and the coded workbook (Summit_workshops_coded_postits.xlsx), which adds the theme, theme area and a "Learning refs" column (L1 to L10) to every post-it, a "Worksheet coding" tab and a "Learning crosscheck" tab mirroring this table. Post-its are referenced as Group #number, matching the transcription workbook.

#	Learning	Figure	What is counted	Where to find it
1	**The focus shifts across the timeframes**	40% (34 of 84)	Short-term post-its in the area Fit, compliance & cost / all short-term post-its	Coded workbook, 'Coded post-its' tab: filter Learning refs contains L1 and Stage (analysis) = Short term (34 rows); remove the L1 filter to see all 84 short-term rows
		64% (34 of 53)	Medium-term post-its in the areas Day-to-day use (17) or Teaching & learning (17) / all medium-term	Coded workbook, 'Coded post-its' tab: Learning refs contains L1, Stage (analysis) = Medium term (34 rows) out of 53 medium-term rows
		41% (14 of 34)	Long-term post-its in the area Impact & evidence / all long-term	Coded workbook, 'Coded post-its' tab: Learning refs contains L1, Stage (analysis) = Long term (14 rows) out of 34 long-term rows
2	**Evidence of impact on pupil learning and outcomes was the most common theme**	13% (25 of 189)	Post-its coded Evidence of impact & outcomes / all included post-its	Coded workbook, 'Coded post-its' tab: Theme = Evidence of impact & outcomes (Learning refs contains L2), 25 rows
		26% (9 of 34)	Long-term Evidence post-its / all long-term	Coded workbook, 'Coded post-its' tab: Learning refs contains L2 and Stage (analysis) = Long term, 9 rows
		One post-it each	Mechanism; theory of change; what would have happened without the tool; which groups it works best for	LbQ Group 4 #1; Tes Group 3 #15; Magma Group 2 #6; LbQ Group 3 #3
3	**Each product drew a different mix of questions**	Tes 24% (6 of 25), 20% (5 of 25)	Tes post-its coded Integration, and Time saving	Coded workbook, 'Coded post-its' tab: Vendor = Tes, Theme = Integration, devices & tech fit (6) / Time saving & workload (5)
		Magma 19% (8 of 43)	Magma post-its coded Pedagogy & teaching practice	Coded workbook, 'Coded post-its' tab: Vendor = Magma, Theme = Pedagogy & teaching practice
		Zero G 21% (8 of 38), 18% (7 of 38)	Zero G post-its coded Data, reporting & insight, and Pupil experience	Coded workbook, 'Coded post-its' tab: Vendor = Zero G, Theme = Data, reporting & insight (8) / Pupil experience & engagement (7)
		LbQ 17% (14 of 83)	LbQ post-its coded Evidence of impact & outcomes	Coded workbook, 'Coded post-its' tab: Vendor = LbQ, Theme = Evidence of impact & outcomes
		LbQ 14 of 15 themes	Number of distinct themes used for LbQ post-its (only Usage & dosage is absent)	Coded workbook, 'Coded post-its' tab: Vendor = LbQ, check distinct values in Theme
4	**Groups questioned how time savings are measured**	6% (12 of 189)	Post-its coded Time saving & workload	Coded workbook, 'Coded post-its' tab: Learning refs contains L4 (12 rows). Original post-its: Tes Group 2 #3, Tes Group 2 #4, Tes Group 3 #12, Tes Group 3 #13, Tes Group 3 #16, LbQ Group 1 #7, LbQ Group 3 #7, LbQ Group 3 #10, LbQ Group 4 #10, Magma Group 3 #6, Magma Group 4 #6, Zero G Unknown Group #12
		6 of 12, from three groups	Workload post-its asking how a saving was quantified, measured or baselined; "hypothetical or tested/piloted" is Tes Group 3 #12	Learning refs = 'L4 (quantification)': LbQ Group 3 #7, LbQ Group 3 #10, Tes Group 2 #4, Tes Group 3 #12, Tes Group 3 #13, Tes Group 3 #16
5	**No group wanted the product company to draw the conclusions**	0 of 7	Groups choosing the product company to formulate conclusions from the data collected and analysed / groups answering that field	Transcription workbook, field 'Data responsibility - Formulates conclusions' in each group's worksheet block. Answered: Tes Group 2, LbQ Group 1, LbQ Group 2, LbQ Group 3, LbQ Group 4, Magma Group 2, Magma Group 4

#	Learning	Figure	What is counted	Where to find it
		LbQ Group 4 crossed out	Product company marked with an X for conclusions	LbQ tab > LBQ GROUP 4 > worksheet row 7, 'Data responsibility - Formulates conclusions': 'Crossed (X mark): Product company'
		71% (5 of 7)	Groups ticking internal team, school leadership or MAT team for conclusions	Tes Group 2, LbQ Group 1, LbQ Group 2, LbQ Group 3, LbQ Group 4
		29% (2 of 7)	Groups ticking only a researcher for conclusions	Magma Group 2, Magma Group 4
		6 worksheets	Conclusions field left blank	Tes Group 1, Tes Group 3, Magma Group 1, Magma Group 3, Zero G Group 1, Zero G Unknown Group
6	**Over half accepted the product company collecting data**	56% (5 of 9)	Groups ticking Product company to collect data / groups answering the collection field	Field 'Data responsibility - Collects data'. Ticked: Tes Group 2, LbQ Group 1, LbQ Group 2, LbQ Group 4, Zero G Unknown Group. Answered: Tes Group 1, Tes Group 2, LbQ Group 1, LbQ Group 2, LbQ Group 3, LbQ Group 4, Magma Group 2, Magma Group 4, Zero G Unknown Group
		25% (2 of 8)	Product company ticked for analysis / groups answering the analysis field	Field 'Data responsibility - Analyses data'. Ticked: LbQ Group 1, LbQ Group 2. Answered: Tes Group 2, LbQ Group 1, LbQ Group 2, LbQ Group 3, LbQ Group 4, Magma Group 2, Magma Group 4, Zero G Unknown Group
		63% (5 of 8)	Internal team ticked for analysis	Tes Group 2, LbQ Group 1, LbQ Group 2, LbQ Group 4, Zero G Unknown Group
		38% (3 of 8)	Researcher ticked for analysis	LbQ Group 1, Magma Group 2, Zero G Unknown Group
		25% (2 of 8)	Independent org ticked for analysis (Magma Group 4 via Other)	LbQ Group 3, Magma Group 4
		Tes Group 1 crossed out	Product company marked with an X for collection	Tes tab > TES GROUP 1 > worksheet row 5, 'Data responsibility - Collects data': 'Crossed (X mark): Product company'. Note Magma Group 4 'Other: school & Product agree on an independent org' is counted as Independent org
7	**Attainment, teacher voice and pupil voice were listed most**	83% (10 of 12)	Data-point lists coded Attainment & progress / worksheets with data points completed	Field 'Data points required for the evaluation'. Coded workbook 'Worksheet coding' tab, column Data points. Attainment: Tes Group 1, Tes Group 3, LbQ Group 1, LbQ Group 2, LbQ Group 3, LbQ Group 4, Magma Group 1, Magma Group 3, Zero G Group 1, Zero G Unknown Group. Completed (12): all worksheets except Magma Group 2 (left blank)
		75% (9 of 12)	Teacher & staff voice	Tes Group 2, Tes Group 3, LbQ Group 1, LbQ Group 2, LbQ Group 3, LbQ Group 4, Magma Group 3, Magma Group 4, Zero G Unknown Group
		67% (8 of 12)	Pupil voice	LbQ Group 1, LbQ Group 2, LbQ Group 3, LbQ Group 4, Magma Group 1, Magma Group 3, Magma Group 4, Zero G Unknown Group
		33% (4 of 12)	Usage & adoption data	LbQ Group 2, LbQ Group 4, Magma Group 3, Magma Group 4
		4% (7 of 189)	Post-its coded Usage & dosage	Coded workbook, 'Coded post-its' tab: Learning refs contains L7: Tes Group 1 #5, Magma Group 1 #8, Magma Group 4 #7, Zero G Group 1 #8, Zero G Unknown Group #13, Zero G Unknown Group #14, Zero G No-Worksheet Group #1
8	**Most planned evaluations sat at the top of the 3 to 6 month window**	Brief: 3 to 6 months	Window set by the session plan (2:40 segment and facilitation notes)	Testbed Design Workshop: Session Plan for Vendors

#	Learning	Figure	What is counted	Where to find it
		15% (2 of 13)	Duration given as a range within the window	LbQ Group 1 (3-6 months), LbQ Group 3 (4-6 months)
		23% (3 of 13)	Duration beyond the window (up to a year)	Tes Group 1 (1 Year), Tes Group 2 (1 year or 2 terms), Magma Group 3 (a year)
		54% (7 of 13)	Research questions with a duration of six months or two terms	Field 'RQ - Duration': Tes Group 3, LbQ Group 2, LbQ Group 4, Magma Group 1, Magma Group 2, Magma Group 4, Zero G Unknown Group. Others: LbQ Group 1 3-6 months, LbQ Group 3 4-6 months, Tes Group 2 1 year or 2 terms, Tes Group 1 1 year, Magma Group 3 a year, Zero G Group 1 blank
		46% (6 of 13)	Pre/post baseline only	Field 'RQ - Baseline / comparison': Tes Group 1, Tes Group 2, LbQ Group 1, LbQ Group 2, LbQ Group 4, Zero G Unknown Group
		38% (5 of 13)	Comparison group (control class, business as usual, non-users)	Tes Group 3, LbQ Group 3, Magma Group 2, Magma Group 3, Magma Group 4. Tes Group 3 counted from 'RQ - Setting / context: Control Group Class'
		8% (1 of 13)	Historical cohort plus pre/post	Magma Group 1 ("prior cohorts + before/after attitudes")
		Single class to several schools	Range of cohorts written in RQ - User group / cohort and Users for this evaluation	e.g. Magma Group 2 "a Year 5 class"; Tes Group 2 "1xSendco + 1xLeader + 1xClass"; Tes Group 1 "Multiple pilot schools"; Magma Group 4 "Students - 120"
9	**Data protection and IT leads were widely included**	78% (7 of 9)	Groups ticking DPO and IT lead / groups answering the evaluation team field	Field 'Evaluation team - who needs to give input'. DPO and IT both ticked: Tes Group 1, Tes Group 2, LbQ Group 1, LbQ Group 2, LbQ Group 3, LbQ Group 4, Magma Group 4. Answered: Tes Group 1, Tes Group 2, Tes Group 3, LbQ Group 1, LbQ Group 2, LbQ Group 3, LbQ Group 4, Magma Group 2, Magma Group 4
		18% (15 of 84)	Short-term post-its coded Integration, devices & tech fit	Coded workbook, 'Coded post-its' tab: Learning refs contains L9, Theme = Integration, devices & tech fit
		8% (7 of 84)	Short-term post-its coded Data protection, safety & wellbeing	Coded workbook, 'Coded post-its' tab: Learning refs contains L9, Theme = Data protection, safety & wellbeing
10	**Participants wanted pupils and parents involved**	90% (9 of 10)	Groups ticking Pupils / groups answering the voices field	Field 'Whose voices and experiences should be included'. Pupils: Tes Group 1, Tes Group 3, LbQ Group 1, LbQ Group 2, LbQ Group 3, LbQ Group 4, Magma Group 2, Magma Group 4, Zero G Unknown Group. Answered: Tes Group 1, Tes Group 2, Tes Group 3, LbQ Group 1, LbQ Group 2, LbQ Group 3, LbQ Group 4, Magma Group 2, Magma Group 4, Zero G Unknown Group
		70% (7 of 10)	Groups ticking Parents	Tes Group 1, Tes Group 2, Tes Group 3, LbQ Group 1, LbQ Group 2, Magma Group 2, Magma Group 4
		31% (4 of 13)	Worksheets writing pupils or parents into the evaluation team / all worksheets	Field 'Evaluation team - who needs to give input', 'Written in Other': LbQ Group 1 ('MAT / PARENTS / STUDENTS'), LbQ Group 2 ('Parents'), LbQ Group 4 ('Pupils', box not ticked), Magma Group 1 ('students')

A2 · Every theme by timeframe, with example post-its

The full list behind the top three themes shown in Section 1. Counts are post-its; the example is quoted as written, with its group and post-it number.

Short term: 84 post-its

Theme	Post-its	What groups asked about	Example post-it (as written)
Integration, devices & tech fit	15 (18%)	How the product connects to MIS, SSO, local authority data and other systems, which devices it runs on, and whether pupils need 1:1 devices.	*"Will IT teams be happy? SSO, MIS integration, DPIA/GDPR compliant"* LbQ Group 2 #8
Evidence of impact & outcomes	9 (11%)	Whether evidence already exists that the product works, for whom and through what mechanism, including requests for RCTs and for schools using it well.	*"What is mechanism by which it impacts pupil outcomes? And what evidence supports that?"* LbQ Group 4 #1
Pupil experience & engagement	8 (10%)	What pupils see and do, and what happens when they finish early, get stuck or go off task.	*"What happens if student keeps saying "I don't know" i.e. task is pitched too high"* Zero G Unknown Group #9
Data protection, safety & wellbeing	7 (8%)	Where data is stored, what is held on children, security, safeguarding and service levels. One group raised the screen-time debate with its board.	*"Where does the data 'go'? How secure is it?"* LbQ Group 4 #5
Inclusion, SEND & differentiation	7 (8%)	Accessibility features and how the product caters for SEND and EAL learners.	*"How does it work for SEND? Is it accessible for all pupils?"* Magma Group 3 #2
Implementation, training & support	6 (7%)	Training and onboarding time, CPD, and how long a pilot would last.	*"How are Teachers onboarded with trainings - length, type, regularity + CPD?"* LbQ Group 3 #2
Purpose, positioning & roadmap	6 (7%)	What problem the product solves and how it differs from competitors.	*"How does your product differ from the other seven pioneers part of the Government programme?"* Zero G No-Worksheet Group #4
Curriculum, policy & context fit	5 (6%)	Alignment with schemes of work, exam specifications, Ofsted expectations and local context.	*"Does this align to Ofsted reporting expectations"* Tes Group 3 #2
Data, reporting & insight	5 (6%)	What reports show, and whether teachers have the data literacy to interpret and act on them.	*"How do you know the teachers have enough data literacy to interpret the figures they are seeing?"* Zero G Group 1 #3
Cost & value	4 (5%)	Price, licence model and cost implications, including outcomes-based contracting. All four were asked of LbQ.	*"What is the cost and licence model"* LbQ Group 1 #2
Features & functionality	4 (5%)	Homework use, bespoke content and question types.	*"Can you set homework?"* LbQ Group 2 #7
Time saving & workload	4 (5%)	Whether the product saves time, including the effort needed for teachers to set work.	*"Will it save time?"* Tes Group 2 #3
AI role, accuracy & safeguards	3 (4%)	The accuracy of AI marking and feedback, and pupils' understanding of AI.	*"How accurate is the feedback from the auto grade?"* Magma Group 4 #4
Usage & dosage	1 (1%)	Whether the product is best used in class or at home.	*"Is this best used in class or @ home?"* Zero G No-Worksheet Group #1

Medium term: 53 post-its

Theme	Post-its	What groups asked about	Example post-it (as written)
Pedagogy & teaching practice	12 (23%)	How the product changes teaching: unintended consequences, deskilling, productive struggle, where it fits in a lesson and which pedagogies emerge.	*"What are the unintended consequvences of this product in the hands of a weak teacher?"* Magma Group 2 #5

Theme	Post-its	What groups asked about	Example post-it (as written)
Data, reporting & insight	7 (13%)	Report options and analytics, parent access, learning maps, class and trust-level insight, and whether teachers act on the data.	*"How will you ensure teachers will use the data and action it?"* Zero G Unknown Group #11
Features & functionality	6 (11%)	Home learning, revision, self-assessment and open-ended questions.	*"Can pupils self-assess?"* LbQ Group 2 #10
Time saving & workload	5 (9%)	How time savings were quantified, and what teachers do with the time freed up.	*"How have you quantified the time saving for teachers/SENDCos.?"* Tes Group 2 #4
Evidence of impact & outcomes	4 (8%)	How the product affects learning, which elements have most evidence, and what would have happened without it.	*"IMPACT: How do you know what would have happeny without the tool."* Magma Group 2 #6
Implementation, training & support	4 (8%)	Ongoing vendor engagement, CPD and school champions, and making sure all staff can use the product whatever their confidence.	*"How do you ensure all staff are equally skilled at using it?"* LbQ Group 2 #12
Pupil experience & engagement	4 (8%)	Pupils' awareness of their progress, their motivation, and knowing when they have practised enough.	*"How do students experience progress?"* Zero G Group 1 #10
Usage & dosage	4 (8%)	Dosage: how often and how much pupils need to use it, and how usage relates to attainment.	*"What "dose" is most impactful & what does it look like?"* Magma Group 4 #7
Integration, devices & tech fit	2 (4%)	Fit with existing infrastructure and device access.	*"Can this be part of our existing infrastructure"* Tes Group 1 #4
Purpose, positioning & roadmap	2 (4%)	How the product has changed over time and how it compares with competitors.	*"How does it compete with Magma?"* LbQ Group 4 #8
Curriculum, policy & context fit	1 (2%)	How often content is refreshed.	*"How often is the platform / questions refreshed?"* LbQ Group 1 #12
Data protection, safety & wellbeing	1 (2%)	Stress-testing the product against the school's data security thresholds.	*"How will you work with us to stress test your product against our system's data security thresholds?"* Magma Group 2 #7
Inclusion, SEND & differentiation	1 (2%)	Whether impact is equal for vulnerable and SEND pupils.	*"Do you see equal impact on children in vulnerable groups / SEND?"* LbQ Group 2 #9

Long term: 34 post-its

Theme	Post-its	What groups asked about	Example post-it (as written)
Evidence of impact & outcomes	9 (26%)	Long-term learning, effects on KS1, MTC and KS2 results, evidence for mathematical thinking, and whether the vendor has a theory of change.	*"What is the effect on long-term learning?"* LbQ Group 4 #17
Data, reporting & insight	5 (15%)	Seeing progress over time, school-level overviews and using data outside the platform.	*"Teacher dashboard: Can a teacher easily see progress over time?"* Zero G Group 1 #11
Purpose, positioning & roadmap	5 (15%)	Development roadmap, scaling strategy, support for more subjects, alternatives and resilience to another lockdown.	*"What is your dev. roadmap?"* LbQ Group 1 #18
Curriculum, policy & context fit	3 (9%)	Adapting to the 2027 curriculum and to primary or KS3 curriculum models.	*"How will you adapt to the new 2027 curriculum?"* LbQ Group 2 #14
Time saving & workload	3 (9%)	Whether the time saving has been tested, and the baseline teacher time. All three came from Tes Group 3.	*"How do you know it saves time = hypothetical or tested/piloted?"* Tes Group 3 #12
AI role, accuracy & safeguards	2 (6%)	Whether AI is used, how it is secured, and whether school data trains models.	*"Is AI included. and will it be used to train models?"* LbQ Group 1 #16

Theme	Post-its	What groups asked about	Example post-it (as written)
Usage & dosage	2 (6%)	Measuring usage and adoption in lessons.	*"Measure of usage of the product."* Magma Group 1 #8
Data protection, safety & wellbeing	1 (3%)	Safety, GDPR and training.	*"Safety. GDPR. Teacher/Student training. Data collection to evaluate."* Zero G No-Worksheet Group #6
Features & functionality	1 (3%)	What the pupil portfolio looks like.	*"What does the student's portfolio look like?"* Magma Group 1 #7
Implementation, training & support	1 (3%)	Long-term training needs, whether training is included in the cost, and roll-out.	*"How much training will staff need? Is it included in the cost now, + 4 yrs. How would you recommend we roll it out? How quickly?"* LbQ Group 3 #13
Integration, devices & tech fit	1 (3%)	Which integrations are needed.	*"What integrations are needed? (MIS, ...)"* Tes Group 3 #11
Pedagogy & teaching practice	1 (3%)	Whether pupils and teachers become reliant on the tool.	*"Do students & teachers become reliant? Can they do math without tools/magma."* Magma Group 3 #14

Not assigned: 18 post-its

Theme	Post-its	What groups asked about	Example post-it (as written)
Features & functionality	5 (28%)	Question mechanics: re-testing, anti-cheating, question versions, home use and tailoring.	*"Can students click off the Qu to use other systems to find the answers? Anticheating?"* LbQ Group 1 #19
Curriculum, policy & context fit	3 (17%)	Fit with different curricula and subjects not yet covered.	*"Can it fit different curricula?"* LbQ Group 4 #23
Evidence of impact & outcomes	3 (17%)	Evidence that learning is happening, and the extent of improvement.	*"Evidence that learning is happening?"* LbQ Group 4 #25
Data, reporting & insight	2 (11%)	Depth of feedback to teachers and cross-subject comparisons.	*"Feedback depth to teachers?"* LbQ Group 4 #28
Implementation, training & support	2 (11%)	Onboarding time and access.	*"Onboarding process, time, and access"* LbQ Group 4 #29
AI role, accuracy & safeguards	1 (6%)	Whether answers are generated automatically.	*"Are answers generated automatically"* LbQ Group 4 #22
Cost & value	1 (6%)	Cost per pupil.	*"Cost per pupil implications?"* LbQ Group 2 #17
Inclusion, SEND & differentiation	1 (6%)	Stretch activities for gifted and talented pupils.	*"Does it have stretch activities for G+T"* LbQ Group 4 #30

A3 · Method, sources and limitations

Data

Handwritten post-its and A1 worksheets from four co-design workshops (one per product) at the EdTech Impact Summit, London, 8 October 2026, transcribed from scans into a workbook with a second, independent reading of every image.

196 post-its were transcribed and 189 kept after removing duplicates and summary post-its (A4). Groups are counted from the post-it sets: 14 sets of post-its and 13 worksheets. The Zero Gravity workshop produced three post-it sets but only two worksheets, so one set of 8 post-its could not be matched to a worksheet.

Company-supplied material: the presentation decks LbQ and Zero Gravity provided before the workshop and presented on the day. These are reported as the vendors presented them and were not verified by EdTech Impact. Tes and Magma Maths summaries have been requested.

Coding

Post-its were kept in the timeframe groups placed them in (short, medium or long term); unsorted post-its are reported as “Not assigned” and the one post-it filed “in between short and medium” is counted as medium term. Each post-it was coded to one primary theme (15 themes) and each theme to one of five areas. Compound post-its were coded by their main thrust.

Worksheet tick boxes are counted as recorded, including crossed-out options. Free-text fields (users, ages, data points, impacts, research question) were coded into categories. Theme, area and category names are analyst labels, not participants’ wording.

Limitations

- Small numbers: 14 groups and 13 worksheets. Percentages are indicative, not statistically representative.
- Coding was done by one analyst and not independently double-coded. The coded workbook allows every decision to be checked.
- Some handwriting was uncertain; the transcription flags these readings for checking.
- Groups had about 20 minutes for the worksheet and were steered to a 3 to 6 month window, so plans show starting instincts.
- Tick-box options were printed on the worksheet, which encourages ticking; write-ins are reported separately.
- Sections left empty are excluded from counts, so each figure states how many groups answered.
- Each workshop covered only one product, so product effects cannot be separated from differences in who was in each room; groups mixed leaders, teachers, researchers and industry experts, and answers are not split by role.
- The LbQ workshop contributed 44% of post-its (83 of 189) and 4 of the 7 groups answering the conclusions question, with 14 of the 63 attendees.
- Attendee numbers per workshop (14 to 17) are given for context and are not used as denominators.
- Vendor claims are available for LbQ and Zero Gravity only; Tes and Magma Maths summaries have been requested.

A4 · Data notes: excluded post-its and worksheet completion

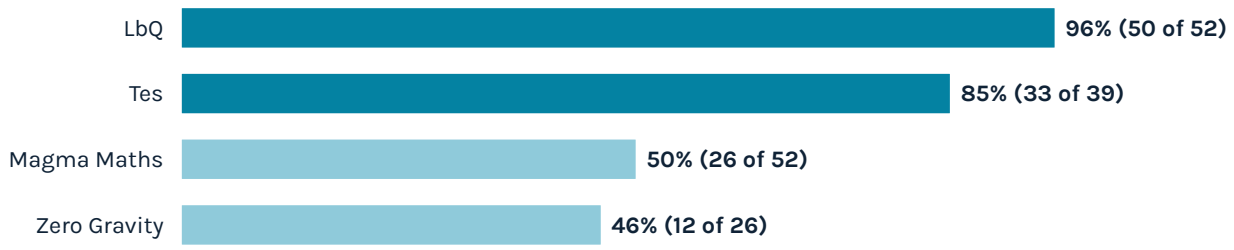
Post-it	Reason for exclusion
LbQ Group 1 #5	Same post-it photographed twice (LbQ_1_Immediate-04/-05)
Magma Group 3 #1	Same post-it as Magma Group 1 #1 (appears in both groups' scans)
Zero G Unknown Group #19	Identical wording to Magma Group 1 post-it; likely mis-filed scan
Zero G Unknown Group #20	Identical wording to Magma Group 1 post-it (Magma portfolio); likely mis-filed scan
Zero G Unknown Group #21	Identical wording to LbQ Group 2 post-it; likely mis-filed scan

Post-it	Reason for exclusion
LbQ Group 4 #21	Summary post-it repeating themes already on other LbQ Group 4 post-its
LbQ Group 4 #32	Summary post-it repeating themes already on other LbQ Group 4 post-its

Worksheet completion

Sections missing from a worksheet were empty, as per the session brief. Magma Groups 1 to 3 completed only some sections, Zero Gravity Group 1 completed few, one Zero Gravity group had no worksheet and one Zero Gravity group number is unknown. Blank sections are left out of the counts, so each figure states how many groups answered.

Figure 18. Worksheet completion by workshop



Completion counts 13 worksheet sections per group: the ten context and evidence-base fields, the two decision boxes, and the research question (counted as complete when at least 5 of its 7 slots were filled). Workshop designers were present at the Tes and LbQ workshops.

Source: workshop worksheets (13 worksheets from 14 groups), EdTech Impact Summit, October 2026

A5 · Theme glossary

Area	Theme	Post-its
Impact & evidence	Evidence of impact & outcomes	25
	Time saving & workload	12
	Usage & dosage	7
Fit, compliance & cost	Integration, devices & tech fit	18
	Curriculum, policy & context fit	12
	Cost & value	5
	Data protection, safety & wellbeing	9
	AI role, accuracy & safeguards	6
Teaching & learning	Pedagogy & teaching practice	13
	Pupil experience & engagement	12
	Inclusion, SEND & differentiation	9
Day-to-day use	Implementation, training & support	13
	Features & functionality	16
	Data, reporting & insight	19
Vendor & roadmap	Purpose, positioning & roadmap	13

A6 · The testbed worksheet

Groups filled in an A1 version of this worksheet after the post-it exercise, focusing on a 3 to 6 month evaluation. It has four sections: Context (types of users, users for the evaluation, pupil ages), Evidence base (data points, who collects and analyses the data and who formulates conclusions, the evaluation team, whose voices to include and how findings are presented), Decision (the impacts that would lead to buying or walking away) and Research question. Shown here is the downloadable version, which adds Product and Duration fields to the version used on the day. The worksheet is turned on its side to show it at the largest size.

