Contents

1. What Does This Report Contain? ................................................................. 4
2. Overview of the Product ........................................................................... 4
3. Executive Summary .................................................................................. 5
4. Detailed Review ....................................................................................... 8
   4.1 Content Quality .................................................................................. 9
   4.2 Pedagogical Alignment ..................................................................... 11
   4.3 Technology and Design .................................................................... 16
Appendix ....................................................................................................... 19
1. What Does This Report Contain?

The section, 'Overview of the Product', provides a brief description of the product and its key features to give the context for the evaluation. The two sections following that present the findings from the evaluation. The Executive Summary provides the overall rating and offers implications in terms of benefits and limitations for teachers and learners. The Detailed Review section provides an in-depth evaluation of the product, categorized under three dimensions (or constructs) – Content Quality, Pedagogical Alignment, and Technology & Design. For each dimension, the product is reviewed on the criteria along with explanations for the rating, and grouped into clusters. Specific examples have been provided in this report to support and elaborate on the evaluation ratings.

The terms, ‘Exemplary’, ‘Valuable’, and ‘Potential to Improve’, used in the report refer to the rating scale for evaluating the product.

- **'Exemplary'** indicates that the product has been designed as per recommended learning theories and research-based evidence.
- **'Valuable'** indicates limited adherence of the product’s design to the recommended learning theories and research-based evidence.
- **'Potential to Improve'** indicates unsatisfactory or lack of adherence of the product’s design to the learning theories and research-based evidence.

2. Overview of the Product

This report evaluated the product Mindspark, Mathematics (Grades 1-2). Mindspark is an adaptive, self-learning platform aimed at providing a unique learning path to each learner based on his or her needs. It is designed using data on learners’ learning gaps and difficulties. It consists of a series of questions, activities, and challenges presented to the learner along with corresponding feedback. Adaptivity is based on learner profile and performance.
Executive Summary

Potential benefits of this product

- Learners can be assured of the correctness of the content and can use the product independently to learn different concepts aligned to their curriculum.
- The product can be used as a great supplement to school and classroom learning as it helps to address common learning gaps for the students.
- Learners at different levels will be able to benefit from the product due to its adaptive nature and presence of features like ‘Effort Mode’ and ‘Challenge Questions’.
- Due to the constructivist approach used for learning, it can also be used as an independent learning resource for practice and to explore and advance beyond the class.
- The learning content is relatable to a good extent for the concerned age group.
- The interface is intuitive to use for learners and Grade 1-2 students will be able to use it independently after a basic introduction to the product.
- Teachers can use Mindspark to track student progress and levels which can help inform their classroom instruction. Also, teachers can create and assign worksheets on the platform directly.

Potential limitations of this product

- The feedback quality is inconsistent across various topics. This might make it difficult for struggling learners to effectively grasp the concept.
- Some Grade 1-2 learners will need help from a teacher or parent to comprehend the language in the questions.
- Learners might be disengaged when too many similar questions are presented upon making a few mistakes. Also, the inability to skip questions or even navigate to a different sub-topic might add to the learner’s frustration. The lack of a conversational or personalized tone in the language of the questions could also lower the engagement of the learners.
- The product doesn’t have sufficient mathematical tools for some topics where they are required. Having these tools would enhance the learning experience and lead to a better understanding of the concepts.
## Executive Summary

### Summary of Review Ratings by Criteria

#### Content Quality: Exemplary

**C1 Content accuracy**  
Not applicable as the product did not contain separate content videos or text; instead “content” in the product was entirely in the form of questions.

**C2 Correctness and clarity in assessment**  
The assessment questions are clear and solutions contain accurate explanations for the majority of the topics.

**C3 Language comprehensibility**  
The intended learner can understand the language with some effort or support.

**C4 Mathematics skill coverage**  
Mathematics mindset and skills required for grade 1-2 (recommended by NEP 2020 and NCF) are covered.

**C5 Curriculum alignment**  
The content is aligned to NCERT and logically sequenced.

**C6 Inclusivity in representation of learners**  
Learners from many sections of the society will find the content relatable but a systematic representation of diverse groups in terms of religion, race and socio-economic is insufficient.

#### Pedagogical Alignment: Valuable

**P1 Constructivist approach**  
The learner has enough opportunities for problem solving and to make sense of the concepts.

**P2 Addressing learning gaps/ alternate conceptions**  
Potential learning gaps have been identified and addressed in some topics but are missing in others.

**P3 Content in context**  
Most of the topics have some real world context but context is not sufficient in some topics to make the learner care about the topic.

**P4 Learner scaffolding**  
The questions are ordered from simple to complex. Hints are present in some learning units which help in simplifying the problems but hints are required in other learning units and should be included.

**P5 Cognitive engagement**  
There is sufficient highlighting and signaling in the product but personalization is missing.

**P6 Motivational features**  
Features like a points system, mastery level and challenge questions keep a learner motivated.
**P8 Learning objective – assessment alignment**
The assessment questions are aligned to the learning objective of the learning unit.

**P9 Pedagogy – assessment method alignment**
Some learning units have pedagogy and assessment aligned to grade-level NEP recommendations, such as scenario-based learning, but this is missing in some learning units.

**P10 Cognitive levels covered**
Sufficient Higher Order Thinking Skills questions are present.

**P11 Feedback Quality**
The feedback is either binary or explanations are not sufficient for the intended learner.

**P13 Adaptivity**
The system adapts and presents questions to a learner based on the learner’s performance and profile.

**P14 Teacher support**
Robust mechanisms to create home works, worksheets and dashboards to track student progress are available.

### Technology and Design: Valuable 😊

**T1 Interface design: Enable intuitive use**
The product is very intuitive to use, and all user actions lead to the expected effect.

**T2 Interface design: Assess consequences of an action**
Some feedback and prompts to reverse errors are present but are not sufficient.

**T3 Learner navigation & pace**
The learners have control over the main chapter selected, although the system decides the learning pathway. Hence, the learners can’t navigate within topics and questions. Also the learner can’t go back to a question.

**T4 Universal Design**
Several Web Content Accessibility Guidelines principles had been followed. However, the product did not have low entry barriers for diverse learners such as learners with varying abilities, learners with special needs, or other characteristics.

**T5 Analytics for learners’ progress**
The dashboard has valuable information for the teachers and learners, but it is actionable to a limited extent.

**T6 Tools to support problem solving**
No mathematical tools to aid in problem solving are present.

**T7 Meaningful interactivity**
Some units have meaningful interactivity features which aid in the understanding of the concept along with basic interactivity.
4. Detailed Review

4.1 Content Quality

Content Accuracy and Clarity .............................................................. 9
Alignment to National Standards ..................................................... 10
Inclusivity in Content Representation ............................................. 10

4.2 Pedagogical Alignment

Learner-Centred Approach ............................................................... 11
Enhancing Learner Experience ....................................................... 12
Assessment of Learning ................................................................. 14
Adaptivity ....................................................................................... 15
Teacher Support ............................................................................. 15

4.3 Technology & Design

User Interface Design ..................................................................... 16
Affordances that Facilitate Learning ............................................. 17
4.1 Content Quality

Content Quality measures the accuracy and content/skill coverage for the grade targeted and the specific domain. This dimension focuses on content accuracy and clarity, alignment to national standards, and inclusivity in content representations.

4.1.1 Content Accuracy and Clarity

Note: This product consisted of a series of questions for the learner and corresponding feedback based on the learner’s response. The “content” in the product was entirely in the form of “questions” and it did not contain separate “content” videos or text. Hence the criterion Content accuracy (C1) is not rated here.

Correctness and clarity in assessment (C2) is rated Exemplary. For most of the topics assessment questions and solutions are factually correct, but in few topics (especially at Grade 1) minor inaccuracies are observed such as in the unit of mass and capacity.

The wording of the assessment questions conveys the intended meaning and clearly states what is expected from the learner. The solutions are complete with appropriate explanations. However at some places, there is the possibility of more solution options being correct, but only one option (most likely one) is chosen in the solution. In some of the long questions, the relevant images are added to support easy comprehension of questions and solutions.

Language Comprehensibility (C3) is rated Valuable: Across many topics, simple, short, and easy-to-understand sentences are used. Even in the case of complex problems the phrasing of sentences is simple, sequenced accurately so that the learner can follow the question as well as solutions. However, there are a few questions where long sentences or abstract wording is used.

Illustrative example:

Starting from 0, a rabbit takes 3 jumps of 2 steps each and reaches 6. This sentence is abstract and difficult to understand for Grade 1 students.

It was also observed (particularly in Grade 2) that in some of the solutions difficult vocabulary and sentence structure is used to explain the answer. This further adds difficulty in understanding.
4.1.2 Alignment to National Standards

Mathematics skill coverage (C4) is rated Exemplary: Math skills required for Grades 1-2 (recommended by NEP and NCF) such as identifying shapes and symmetry, comparisons and classification along one dimension at a time, etc. are covered comprehensively.

Curriculum alignment (C5) is rated Exemplary: The content is covered extensively and topics are chunked appropriately into small subtopics and the conceptual understanding of the topics is progressively built through a series of questions. It is observed that broadly, all the topics and sub-topics covered were aligned with the content present in NCERT textbooks for Grades 1-2. However, the chapter names do not exactly match, which might make it slightly difficult to navigate for a child trying to practice a particular chapter, which he/she would have done at school.

4.1.3 Inclusivity in Content Representation

Inclusivity in the representation of learners (C6) is rated Valuable: The content includes gender diversity as the names used in the questions show a balanced representation of boys and girls. Similarly, the content also covers different socioeconomic backgrounds, with the objects such as matka, swimming pool, car, bicycle, birthday caps, water tank, etc. are part of examples. Thus, some effort has been made to include diversity in various components of the learning content.

The product does not specifically include examples pointing towards a particular religion or culture, so the product is fair and unbiased in terms of addressing the diversity of religion. However, it was observed that in a few places the graphics describing these characters are mostly represented as fair-skinned in a setting with a particular socio-economic background (urbanized and modern home settings) which might create a bias for the learners. (Mostly observed in Grade 1. Kids are very impressionable at that age).

Illustrative example:
The apple is heavier than the leaf because the balance is leaning towards the pan with the apple.
4.2 Pedagogical Alignment

Pedagogical Alignment includes criteria focused on adaptivity, learner-centred pedagogy, enhancing learner experience, assessment of learning, and teacher support. It measures the extent of alignment of the pedagogical strategies with national educational policies, Learning Sciences theories and design principles to create a meaningful learning experience.

4.2.1 Learner-Centred Approach

<table>
<thead>
<tr>
<th>Constructivist approach</th>
<th>Addressing learning gaps</th>
</tr>
</thead>
<tbody>
<tr>
<td>(P1)</td>
<td>(P2)</td>
</tr>
</tbody>
</table>

Constructivist approach in pedagogy design (P1) is rated Exemplary: Many topics follow a constructivist approach through questions and various activities. Those activities allow learners to connect the content with what they know. There are a very well-sequenced set of questions on a topic that progressively addresses the concept from multiple angles or at different levels of complexity.

**Illustrative example:**

When learning about “Comparing weights”, the questions progress from: Feather vs rock to a Large feather vs a small rock to Books vs Watermelons to Comparison of multiple objects at a time (5 feathers vs 2 rocks) Empty bottle vs a filled bottle of water.

**Illustrative example:**

Another example is where questions are generally posed in multiple ways to cater to an understanding of various aspects of the topic, and various opportunities are provided to apply the understanding. For example, there is a subtopic ('Addition facts for 10') where the learners try to make 10 using different numbers. Some ways in which this is presented are:

- Which of these sums would make 10? (A playful activity is also present for the same)
- What must be added to say, 3, to make 7? Again, a good activity is present to support this kind of problem. (Using + and - buttons to increase/decrease the number of objects).
- Picking the correct 'match the following image' where all numbers are matched accurately to get 10.

In some of the topics, learners are allowed to interact with the content through hands-on activities where they try out various permutations and check the result. For example, in the concept of counting (Grade 1), some objects are shown on the screen and the user is required to press the buttons matching with the number of objects. If an incorrect response is made, the system buzzes and gives the learner another chance.
Addressing learning gaps (P2) is rated Valuable: In most of the topics, the learning gaps are addressed through various examples.

Illustrative example:

While comparing the weight of different objects, the misconception of ‘big might be heavier’, or ‘more objects might be heavier’ has been well addressed. Another example is in multiplication, a correct representation and vocabulary are used always between (6x2) vs (2x6). [6 times 2 vs 2 times 6, 6 groups of 2 objects vs 2 groups of 6 objects.

However, there are some instances where there’s a possibility of learners forming misconceptions due to the way some ideas are presented.

Illustrative example:

In the ‘Concept of Time,’ (Grade 1), there are questions such as, 'Which activity would take more time'? [Walking to your school vs eating your breakfast]. Having an explicit correct answer for such questions might lead to confusion.

Another example is in addition (Grade 2) where it’s asked which is a faster way to add on fingers. (6+2 vs 2+6), where neither is definitely right or wrong. The feedback to such questions should be suggestive rather than having a definite correct answer.

4.2.2. Enhancing learner experience

| Content in context (P3) | Learner scaffolding (P4) | Cognitive Engagement (P5) | Motivational Features (P6) |

Content in Context (P3) is rated Valuable: Most of the topics have some real-world context for learners (Grade 1-2) in terms of the objects used. Some of these include stationery items, toys, sweets, furniture items, leaves, trees, people, and animals which a learner could relate to. Many topics have numerous examples from a day to day life, like 3D shapes using a variety of objects found at home. ‘Spatial understanding’ (Grade 1) presents multiple examples showing objects placed relative to each other. Some topics have daily activities which a learner could relate to, like going to school or brushing their teeth.

However, it is observed in many topics that the objects used are not placed in scenarios that will make a child really want to care for the concept or the topic, and make connections to why the concept is useful.
Illustrative example:

In the unit of addition, questions are based on the addition of objects, but no examples and questions of situations or scenarios are shown where adding objects is useful/needed. Some possible examples could be counting a sum of money at a shop or the total number of toys or sweets that you have.

Learner Scaffolding (P4) is rated Valuable: Questions are presented in a way that helps the learner to form a correct mental model of the concept in most cases.

Illustrative example:

Animation is used to introduce addition, where, say, (6+1) is shown as 6 balls in a dotted box and 1 ball in another dotted box, which eventually are placed into a single box.

The content across the topics as well as within the sub-topics is well scaffold and ordered logically helping learners to gradually build an understanding of the topic. Some interactivities also present multiple levels of increasing difficulty, such as the gamified activities at the end in the topics of addition and counting (Grade 1).

However, explicit scaffolds like hints or prompts which would enable a learner to figure out the answers without jumping to the complete solution and answer are missing in most topics for this Grade range. Also, there is a lack of support for a struggling learner to come out of a stuck situation. Though an effort mode is present, it doesn’t always lead to the very specific kind of help that the learner might require.

Cognitive Engagement (P5) is rated Valuable: In this product there are many visual cues provided such as highlighting of the important text (making it bold or different colors), having a border on some aspects of the image, or using dashed lines where appropriate. But personalized conversational style is not seen in any topic and formal book language is used for question and solution presentation. Questions are still in the third person and don’t necessarily invite the learner to engage deeply with the problem.

Motivational features (P6) is rated Exemplary: The product has enough features to make sure the learners are motivated in their learning journey. There is a points system (Sparkies), which a learner earns upon high accuracy as well as completing a unit. If the learner answers 3 questions correctly in a row (consecutively), they collect one Sparkie. These Sparkies feed into a leaderboard announcing the names of learners at various levels (class, standard, city, national level) etc., thus providing some competitive motivation.
4.2.3 Assessment of Learning

Learning objective – assessment alignment (P8) is rated Exemplary. The questions in each topic and concept are very well aligned with the learning objective.

Pedagogy-assessment method alignment (P9) is rated Valuable: The assessment pedagogy method recommended for this grade range is play-based, which the product caters to some extent with the help of interactive games in some topics, but the nature of the product is such that most of the assessment is direct and explicit question-based rather than an implicit assessment based on playful activities.

Cognitive levels covered (P10) is rated Exemplary: The questions are present at various cognitive levels, ranging from recall and understand to apply and create. The majority of the questions are at an 'apply' level of Bloom's Taxonomy.

Feedback Quality (P11) is rated Potential to improve: The feedback quality needs a lot of improvement even though there are some instances where the explanations are good. It’s been repeatedly observed in various topics that quite a few questions in the beginning just contain binary answers without any explanations.

Illustrative example:

In the concept of counting, where a learner needs to count the number of objects shown, the explanation just contains the answer. Ideally, the objects could have been numbered or the counting process could be shown as a simple animation.

Such instances are seen a lot in many other topics like addition and pattern identification. There are questions where an incomplete pattern is shown which needs to be completed. It would be valuable to have explanations that show how the missing part fits into the complete image.

Also, both options could be possibly correct in certain questions. For example: “Which is a quicker way to add? In such cases, explanations need to be more suggestive than final.
4.2.4. Adaptivity

Adaptivity (P13) is rated Exemplary: The learning path in the product adapts according to the performance of the learner. Questions of various difficulty levels are present which are presented to the learner based on the performance. The questions gradually vary in various respects as the learner continues giving a wrong answer.

Illustrative example:
Following is an example from the concept of multiplication. A question (showing multiplication as repeated addition) like \((7+7+7+7+7)\), if answered wrong, would lead to questions of lower difficulty in terms of smaller numbers to lesser numbers to going back to counting using groups of objects with images.

An explicit effort mode is present, which shows up in case of too many wrong attempts. The effort mode has easier questions. However, it is observed in some instances that the effort mode contains a set of repetitive questions which the learner had answered correctly earlier. On the other hand, when the learner is performing with high accuracy, increasingly challenging questions are posed.

The product also takes some initial information (like the grade, age) from the student and does basic adaptation by showing relevant content for that grade.

However, no evidence is found of adaptivity based on the learner’s interaction with the system. For example, if questions are answered at an unreasonable pace, there is no check or a prompt. There aren’t any user interaction patterns (like always entering a particular option with thinking) which the system tries to identify and adapt accordingly. However, since adaptivity based on performance and profile is present, the product is rated Exemplary.

4.2.5 Teacher Support

Teacher support (P14) is rated Exemplary. Appropriate teacher support is available within the product. Robust mechanisms are available to create homework, worksheets, and dashboards to track student progress. Teachers can create and customize their own questions or use textbook questions. They can choose individualized topics for students, curate personalized student journeys plus track class performance at an aggregate level.
4.3 Technology and Design

Technology & Design measures how well the technological affordances and the user interface design integrate with the pedagogy and context to promote a meaningful learning experience for all learners. The criteria in this dimension focus on user interface design and affordances that facilitate learning.

4.3.1 User Interface Design

| Interface design (T1 and T2) | Learner navigation and pace (T3) | Universal design (T4) |

Interface design: Enable intuitive use (T1) is rated Exemplary. The product is very intuitive to use, and all user events lead to the expected effect. All buttons are clearly visible, with text or other cues (like a tick mark on the submit button), and are placed in meaningful locations on the screen. Only meaningful actionable items are present on the interface and it doesn’t contain any distracting elements.

Interface design: Assess consequences of an action (T2) is rated Valuable. The product misses out on a few important aspects regarding conveying the consequences of an action. There isn’t much feedback available, for example, when the user progresses to a new subtopic within a learning unit. While the learner is solving a question, a simple prompt should be shown indicating the progress. Also, the error handling is not effective as there are no warnings or prompts when the user performs certain actions like ‘Logout’ or ‘Done with this topic’. The system prompts ‘You’ve gone idle’ if the learner gets inactive. In the unit of Addition and Subtraction - if the question requires the learner to perform ‘Subtraction’ but performs ‘Addition’ by mistake, the system prompts them to “Did you add? Check again?” However, such opportunities to reverse the actions which might have happened by mistake are very few.

Learner Navigation and pace (T3) is rated Valuable.: Learner navigation is important so that learners can move across different topics as and when they wish, and can work on a particular topic whenever they want. Also, the pace at which the learner wants to attempt the questions is completely flexible, apart from some timed activities which are intended to make some aspects playful. However, it is not possible to switch between subtopics, go to the desired interactivity (which is usually present at the end of a subtopic), go back to the previous question, or skip a certain set of questions. But for each topic, the learner does get to see all the questions attempted so far in a section called ‘How I Did’, which helps him/her revisit the entire set of questions along with the solutions.
Universal Design (T4) is rated Valuable: Since the product also caters to a few very simple aspects of universal design. Concerning operability, some key actions can be performed from the keyboard (without needing a mouse), like submitting an answer and moving to the next question. Many expected and important aspects are missing which would help the content accessible to a diverse set of learners. This includes having a text alternative or captions for images, a voiceover option for reading questions (which is especially essential for this grade range), or voice input for sharing answers. Some interactivities do have a voiceover explaining the task, but such instances are rare.

4.3.2 Affordances that Facilitate Learning

| Analytics for learners’ progress (T5) | Tools to support problem-solving (T6) | Meaningful Interactivity (T7) |

Analytics for learners’ progress (T5) is rated Valuable: The product has a dashboard that is useful to the teacher in terms of the information it presents, which can help guide instruction in the classroom. It shows each learner’s learning trail (the topics and questions which they solved), as well as the overall performance on different topics. The teacher can also see the overall class performance on various topics and subtopics. However, it does not provide sufficient guidance to the teacher on identifying where the learner needs to put in the effort or where additional support should be provided to make progress. The dashboard doesn’t allow the teacher to assign different topics or questions or activities to specific sets of students based on their performance or interaction with the system. There is, however, a possibility to activate or deactivate a particular topic for the entire class. A dashboard for the learner is present as well, where the student can see the questions they have attempted so far, and how well they did in each subtopic. However, as observed with the teacher dashboard, the actionable aspect is missing. A possibility for the same could be to have actionable buttons based on learner’s performance on different topics.

Tools to support Problem-solving (T6) is rated Potential to Improve: No specific Math-related tools (like rulers, clocks, etc.) have been found for this Grade range.

Meaningful Interactivity (T7) is rated Valuable: The interactivity affordance of the product is of good quality. In many places when required and possible, there are meaningful activities where the learner can interact with the system. Also, more interactive activities have been observed in Grade 1 as compared to Grade 2. They are implemented using different ways like text inputs, drag-and-drop, clicking on the screen (say, to burst balloons), buttons to increase and decrease quantities, radio buttons, and drop-down menus.
Illustrative examples:

- A good example of interactivity is in the concept of addition where the learner needs to make a total of 10 and can add or remove objects to a basket using + and - buttons.
- A brilliant example is also in the unit of counting, where the learner presses a button as many times as the number shown on the screen. The system buzzes and gives the user another chance if it’s incorrect. There is also evidence of drag-and-drop features, for example: dragging pictures showing different events to arrange them in a sequence (based on time).

However, some concepts lack interactivity features where they could be really useful. An example is the ‘basic understanding of weight’ where a weighing balance is used. A drag-and-drop feature that could allow learners to add different objects on both sides of the balance will help in making sense of the concept. Also it has been noticed on multiple occasions that a drop-down menu is used where there are only 2 options to choose from. In such cases, it is much better to use radio buttons. In some places where drag-and-drop is used, learners can be allowed to select multiple items and drag them all together. This is especially helpful when too many objects need to be dragged, as observed in the topic of counting.
How does the EdTech Tulna evaluation work?

FRAMEWORKS

EdTech Tulna frameworks define a set of standards for quality design of EdTech products. A rigorous and research backed process is established and applied for the creation of various nuanced frameworks. These frameworks are use-case specific to enable transparent and precise, high stakes decision making. The process includes considering existing research literature, feedback from the ground on multiple stakeholder needs and an appreciation for the quality of solutions currently supplied in the ecosystem.

The frameworks are categorized along the three dimensions of Content Quality, Pedagogical Alignment, and Technology & Design to capture a holistic view of the quality of the product design. The frameworks are also made available at varying levels of depth for varying stakeholder needs and range from supporting governments and institutions in making high stakes, rank based, adoption decisions, to providing a brief overview of the key criteria to be considered while designing a product.

TOOLS

Each Tulna framework is accompanied by a toolkit that is specifically designed to guide experts to evaluate EdTech products. These toolkits are customized to the type of EdTech solution, grades, subjects, to drive meaningful and nuanced evaluations. The tools are informed by research as well as iterative empirical study and tested for inter-rater reliability and validity. A typical toolkit consists of rubrics and reviewer guidelines to enable evaluators to interpret the framework and conduct unbiased evaluations. Each criterion within the framework is rated along a three-point rating scale - ‘Exemplary’, ‘Valuable’, and ‘Potential to Improve’ - indicating the level of alignment with expectations laid out in the framework. Toolkits include supporting materials - videos, templates, and example illustrations - to guide experts while conducting evaluations.

PROCESS

Each product goes through a rigorous review process that takes approximately 160 hours for four grade ranges K-2, 3-5, 6-8, and 9-10. Each review team is designed to be independent and neutral. A typical expert review team consists of 3-4 members who are subject matter experts, instructional designers, user-interaction experts, user-experience design experts, and professionals with experience in teaching and implementing EdTech in field settings. Each review team has an anchor of at least one experienced evaluator.

Each member of the expert review team undertakes a two-week long intensive training on understanding the frameworks and the subsequent application of its toolkits to conduct evaluations. For each product, the review team applies a systematic sampling strategy and decides the representative learning units that will be reviewed. The team collectively reviews a subset of the learning units to check for convergence and establish inter-rater reliability. Team members then individually review the remaining learning units. The team finally meets to synthesize key points and takeaways of each review and elaborates their reviews into an in-depth report, which is overseen by the experienced evaluator.

The role of the product company is limited to an initial demo which supports the review team to deepen their appreciation of the intended use of the product, and its scope. The product company is then provided the final reviews and their unedited responses are published alongside the expert evaluations on the Tulna evaluation center.