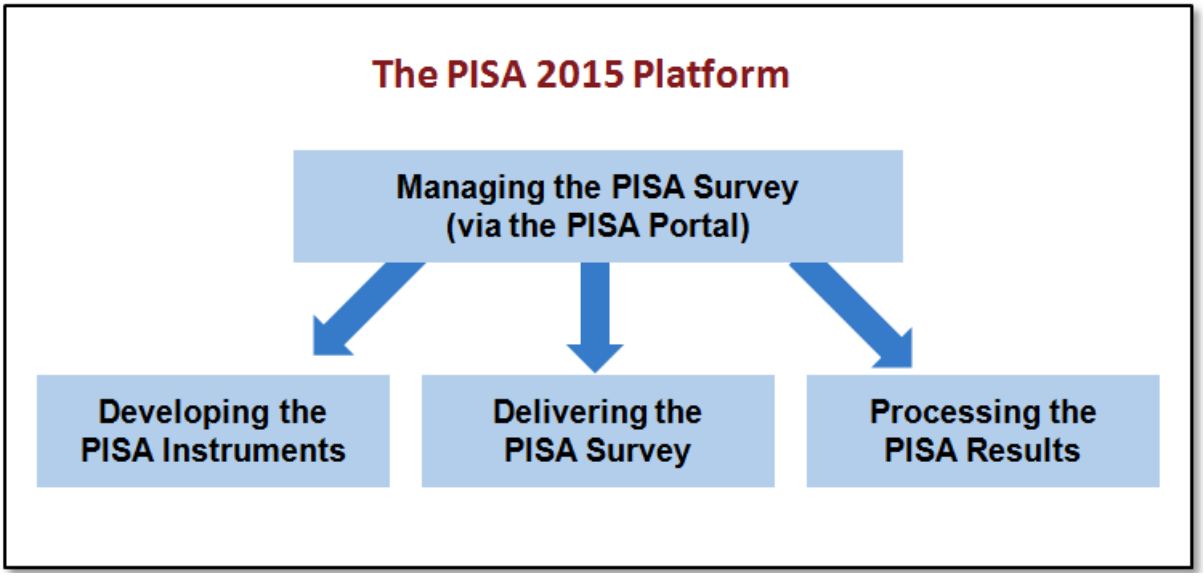
**PISA 2015 COMPUTER PLATFORM DEVELOPMENT**

1. As currently planned, the PISA 2015 platform will operate in four key areas of the PISA project, as shown in Figure 1. The platform will provide access to the tools and processes needed for the overall management of the survey, for building the survey instruments, delivering the survey, and processing the survey results. Note that the design for the PISA 2015 platform and its associated functionality will be finalized by the summer of 2012 and thus the plans presented here will be refined to reflect ongoing work with the OECD and other Core contractors.

**Figure 1**. The Role of the PISA Platform




**Managing the PISA Survey**

2. The primary objective of the PISA Portal is to provide a structured and easy-to-use access point to the key resources needed by the various stakeholder groups involved in the development and delivery of PISA 2015. To this end, we have proposed a role-based model for building the underlying architecture of the portal, focusing on the various roles of the project participants, the functions to be performed by individuals in each of those roles, and the tools they will need to fulfil those functions.

3. There will be a set of common functions needed by all the Core contractors and other stakeholder groups involved with the survey. While specifying the common set of functions will be an evolving effort as PISA contractors join the team, some of the tools and sets of information likely to support those functions include: a PISA project timeline, information updates, communication tools, and a content management system.

4. The development and conduct of PISA will also involve individuals operating across a range of different roles within the project. To meet their needs, the PISA Portal will also be organized around role-specific functions, providing the means for connecting these individuals with the tools and information they need to design and deliver PISA 2015. Each key role in the project will be

associated with a number of different activities and resources related to those activities. The portal will provide tools and information to support role-specific functions in PISA, some of which are listed below.

- *Task-dedicated tools.* These tools will be the primary focus of the PISA Portal and will be organised around three major tasks of the PISA survey: developing the PISA instruments, delivering the survey, and processing the results. These tools and associated information will be made accessible on the portal at appropriate points in the project, orchestrated by the PISA workflow engine (described in greater detail below).

- *User management and profiling functionalities.* These functionalities will make it possible to register participants working on the survey and to assign roles in the PISA workflow. The portal will contain customised screens, or work environments, for each type of user. On that screen, users will see the set of applications appropriate to their project role.

- *Direct data and resource access.* Access to data and resources will be controlled by one or more Core contractors and will be provided to users through dedicated widgets made available at various stages of the PISA survey. Widgets are small software packages that support the integration of a wide range of functionality on a single web page. In addition to resource-dedicated applications such as authoring and previewing tools, widgets enabling downloads and uploads of resources such as translation files, items and response data will be activated for selected roles.

- *Task monitoring facilities.* These can be offered to all project participants either as a reminder of their currently assigned tasks and the status of those tasks or, for users in management roles, to allow them to follow the status of a set of running processes. For example, this type of facility will allow NPMs to monitor the progress of the different tasks for which they are responsible and have a real-time accounting of the project status in their country.

5. The basic delivery for the PISA Portal service will be via online Internet access through a web browser. As a Web 2.0 application, the PISA Portal will be usable on any type of last-generation Internet browser. At present time, the portal is compliant with the latest versions of Mozilla Firefox, Microsoft Internet Explorer, Apple Safari, Google Chrome and Opera. Thanks to the Unicode encoding, alphabets such as Chinese, Arabic or Hebrew that contain special symbols are supported.

***Workflow Management***

6. The success of complex computer-based assessment projects does not reside exclusively in the technical platform challenges but also in the definition of organisational processes and the monitoring and reporting of those processes. For PISA 2015, there will be processes and procedures that will be the responsibility of the Core contractors (such as test development and verification) and others that will be the responsibility of the NPMs and National Centers (such as item submission, review and translation). TAO provides a workflow management system that supports those organisational processes and provides the tools to drive them. The TAO platform (as distributed on the official web site http://www.tao.lu) as well as its PISA specific deployment and customisation includes workflow engines that will be used to drive all PISA survey processes. This workflow system will be accessible through the PISA Portal to all involved in the work associated with PISA 2015, providing the access point for every user to perform his or her IT supported functions in the PISA project.

7. Orchestrating the survey processes within the PISA Portal will consist of defining these processes as sequences of activities and assigning them to the roles associated with those activities. A sophisticated workflow engine, implemented in the TAO platform and released under the GPL v2 licence, drives those processes and presents the right TAO functionality to the right person at the right moment. For users in a given role, an instance of their work process will be initiated and they will be notified to connect to the portal and perform the assigned activities. Once the task is completed, the

next actor will be notified and receive the necessary information. This process can be automated or, where not possible, the system will at a minimum be able to identify the next person responsible for each activity in a workflow, allowing managers to keep work processes functioning smoothly. Specification of the workflow, associated role(s), and activity definition is flexible and can be modified by the contractor managing a specific aspect of the project to accommodate changes in processes or meet the specific needs of a given country or user group. As a corollary, process reallocation from one user to another is possible, for instance, to transfer the work from one translator to another should someone withdraw from the project. A series of monitoring tools will make it possible to follow progress on each process based on a specified timeline and to take actions such as re-assigning an activity/operation, aborting or modifying a process, or alerting a user based on that information.

**Developing the PISA instruments**

8. Developing the complex array of materials for PISA 2015 requires a platform that has broad, well-established support for standard computer-based item types in addition to the extensibility necessary to add new, innovative items that expand the measurement of the PISA domains. Additionally, the platform must provide tools that allow test developers to efficiently and easily manage a broad range of content types and to track the status of this content through its entire lifespan.

***Item Development Process***

9. The item development process for all computer-based items in the PISA 2015 assessment – those in science, reading, math and collaborative problem solving, and possibly financial literacy – will include the same set of sequenced tasks. The role of the platform will be to support each of those tasks.

10. One function of the PISA Portal will be to provide the tools and information needed to manage the item development workflow. Additionally, the portal will contain resources for each of these processes:

- Storyboarding and prototyping new items,
- Authoring both trend and new items,
- Making items available for review by subject matter experts, participating countries, other Core contractors, and the PISA Governing Board as well as providing a repository for collecting these reviews,
- Translating and adapting selected items,
- Verifying items and refining translations/adaptations as needed, and
- Assembling the assessment instruments so they can then move from development to delivery and, finally, processing of the results.

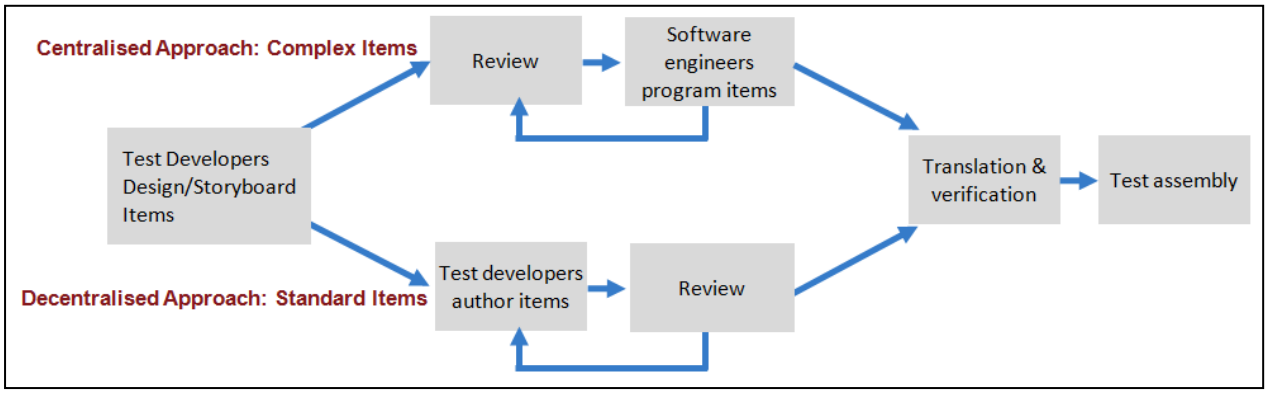
***A Hybrid Approach to Item Authoring***

11. The platform will support a range of item types and response modes which we have classified as "standard" item types and "complex" item types. Supported standard item types include: multiple choice (both single selection and multiple selection), ordering, association and matching, fill in the blank (dragging text or image), clicking on text or graphic image, positioning an object, using a slider to select a number between an upper and lower bound, and drawing. String entry and text entry are also supported, although automatic scoring is unlikely for extended text entry.

12. The TAO platform also has the capability to develop complex item types capable of assessing aspects of the cognitive constructs that cannot be measured in paper-based assessments. Examples include simulation tasks where students can collect and analyse data as well as dynamic problems where information or features of the environment change as students interact with the material. In addition, the platform can support simulations of natural phenomena, the inclusion of digital tools for managing and displaying data, and adaptive routing through complex scenarios. The development of functionalities not currently supported by the platform will be negotiated with other Core contractors to ensure that PISA 2015 can be successfully delivered within the constraints of the project scope and timeline.

13. As shown in Figure 2, the platform will support a hybrid approach to item development where we provide decentralised authoring tools for item writers to develop standard item types but also allow centralised authoring for more innovative, and therefore complex, item types.

**Figure 2**. Hybrid Development Model




14. For simple and standardised items, the platform will provide an authoring tool for the test developer. Within the constraints of the tool functionality and the item types, the developer will have the freedom to concentrate on content, rather than struggle with technical aspects of the item. With simpler items, the transformation of the design from the authoring tool to the executable form is straightforward, enabling the test developer to view and revise the item in an efficient trial-and-error process that will result in a near-final version when the completed item in posted. By implementing the existing TAO authoring tool for PISA, this decentralised approach can be used for a range of item types including multiple-choice, fill-in-the-blank, matching, ordering, text entry and selecting a word or hotspot and employing response modes such as clicking, typing and dragging information on the screen.

15. The kind of authoring tool required to allow decentralised authoring of more innovative item types, including simulations and items with complex interactions or specialised functionality, is necessarily complex. Building such a tool would require a considerable investment of time and resources. This is time which will be better spent developing the underlying platform functionalities to support innovative item types. The additional drawback to this kind of complex authoring tool is that it would require test developers to invest a significant amount of time in both learning how to use the tool and authoring their items. The implementation of these innovative item types will therefore require the expertise of Core 3 IT specialists and software engineers. Taking a centralised approach to the authoring of these items is much more likely to allow test developers to create new and innovative types of items – an important goal for PISA 2015. Adopting a centralised approach for innovative item types will allow test developers to focus their development efforts on conceptualising and designing these items.

16. Under the centralised approach, test developers will be responsible for developing item content, specifying how interactions should work, and defining associated scoring and data capture requirements. This process will occur using storyboards that will include a description of the

scenario, the task(s) associated with that scenario, the content and interactions within the complex stimulus, and the scoring rules (in the case of automatic scoring) or the scoring instructions (in the case of human scoring). The platform will support an approach to storyboarding that will allow test developers to use their tools of choice (such as a Word document or set of PowerPoint slides) for laying out the content and structure of an innovative task. Storyboard information will be stored in the item repository, along with metadata that links tasks to the frameworks in a way that will allow test developers and others to search and organize that information.

17. The platform will support complex item types that can be implemented as Open Web Items (OWI). Open Web Items (http://forge.tao.lu/projects/tao/wiki/Free_Form_Items) are based on standard Web technologies – XHTML, CSS and JavaScript – and allow programmers to implement items that are highly complex and include custom interactions. Complex and innovative items can be programmed as web applications with the appropriate web development tools. Once this is done, they can be connected to the TAO platform using the Item Runtime and Scoring APIs.

18. Two existing algorithms that Core 2 proposes to support are Dynax and Finix. Dynax is a linear structural equation (LSE) system solver. In LSE-type contexts, the animation of an item is reduced to a set of rules taking input parameters (resulting from student interactions) and returning output values (impacting the graphical user interface of the item); output values also may be reused as input parameters. In the technical implementation of Dynax, the rules are transformed into equations, the parameters that are modifiable by the student are mapped to exogenous variables, and the internal parameters correspond to endogenous variables. A student is expected to solve the presented problem based on the understanding gained through the manipulation of the underlying system.

19. While it is desirable that the complex items built for PISA 2015 provide a range of content coverage and interaction types, it will not be practical to implement each from scratch. Core 2 will provide a set of core algorithms that can be leveraged across a variety of scenarios and simulations to support standard control strategies and operations. These core algorithms will allow the Core 3 programmers developing these complex items to rapidly construct rich environments using well-tested and proven code bases.

***The Item Repository***

20. The core of the systems for building the PISA assessment is the item repository where all cognitive items, both trend and new, paper and computer, will be housed. The repository will contain all versions of each item, including the international master and national versions as used in both the Field Test and Main Study. If available from the incumbent PISA contractor, views of trend items from previous PISA cycles (including PDF versions or scanned images) could be included as well. The Core 3 contractor will be responsible for populating the repository with both old and new items – importing items used in previous PISA cycles and developing the new items for PISA 2015. We plan to support those efforts by providing tools, training and other resources and to work collaboratively with Core 3 and Core 6 staff to define the resources needed to build a robust item repository that will support PISA 2015.

21. Beyond the items themselves, the item repository will contain a wealth of information including item classifications, history and statistics. This information will include scoring rubrics, coding information identifying the framework components assessed by the item, translation and adaptation guidelines, layout and formatting requirements, statistical information, reviewer comments, and notes regarding any country-specific issues related to each item. In addition, the repository will contain all questions for the background and school questionnaires and their associated information. Beginning with the task of importing trend items and their associated information, the item repository will grow as new items are designed and authored and will continue to evolve over the course of developing and delivering the Field Test and Main Study instruments and processing those data.

22. The repository will be much more than a simple storage container for all the content associated with the PISA survey. TAO supports a very rich and flexible item repository. Because it is built on Generis®, a system for modelling and managing knowledge, the repository can easily be shaped to support a wide range of data and relationships. It will contain tools allowing users to leverage the various sources of information in the repository. Users will be able to manipulate content by previewing, sorting and filtering information. For example, a test developer working to select trend items for the science assessment could use a search function to generate a list of all the previous science items coded as "demonstrating knowledge and understanding," associate statistical information with that list including data about the difficulty of each item and how well it performed, filter out any items performing below a specified criteria, and then generate a summary report. This ability to search and filter based on item metadata will support a powerful test assembly function. Manipulation tools will make it possible for information in the item repository to be selected, organised and presented in a variety of ways, increasing its usability and utility across the variety of different user types.

***Supporting Item Translation***

23. The PISA platform must support the translation process with tools that translators can use to translate assessment materials, that verifiers can use to review and comment on these translations, and that test developers and countries can use to view and approve the final translated items. TAO uses two key technologies to enable this process: an open-source tool for translating digital text and the TAO workflow engine. TAO supports working in both online and offline modes. When working online, TAO provides live previews of the translated items. This important feature makes it possible to identify and correct any layout and context issues as early as possible in the development process.

24. The TAO workflow engine will support translation of items and PISA questionnaires through its web-based workflow portal. With this workflow engine, the translation, review and approval processes can be defined, step by step. Because the engine is flexible, we will be able to support a variety of approaches to translation, including standard back translation as well as the double translation with reconciliation process currently used in PISA and PIAAC. Once the Core 3 contractor defines the approach for PISA 2015, we will be able to implement the appropriate processes and tools within the PISA platform.

***Paper and Pencil Booklets***

25. In terms of the paper booklets, the primary responsibility for assembling final, print-ready versions for PISA 2015 rests with the Core 3 contractor. In keeping with the proposed approach for the computer-based items, the TAO workflow engine can be used to manage and coordinate the development process for the paper-based items. The PISA Portal will contain tools allowing managers to schedule and assign tasks associated with the paper-based process. As previously noted, the Core 3 contractor will be able to specify the workflow processes and associate roles and activities with those processes.

26. In collaboration with the Core 3 contractor, we will work to develop templates for including the necessary metadata to be associated with each item. Information such as scoring guides, translation and adaptation notes, history of use, framework coding (both based on the 2015 frameworks and previous frameworks) and statistical data can be attached to items, along with reviews and comments regarding their suitability for inclusion in 2015. As we have noted, users will be able to manipulate this content by previewing, sorting and filtering information in the repository. Once the Core 3 contractor has completed the selection, layout and assembly process for the final paper booklets, we expect that printable PDF versions of the assembled tests will be stored in the item repository so they can be selected and exported in ZIP archives for printing. We expect that access to these materials and the definition of this final stage in the process will likely be determined collaboratively between the Core 3 and Core 4 contractors.

**Delivering the PISA survey**

27. The platform will allow flexible delivery options to address the variety of school infrastructures and avoid the reliability issues associated with delivery over the Internet.

- Offline option: Memory sticks are used for deployment and data collection. This option has been reliably used in PISA 2012.
- Semi-online delivery option: testing software is downloaded to the school server or testing computer and once the test is complete, results are uploaded to the national centre. This option is suitable for schools with robust Internet connectivity.

28. Probably the most reliable deployment strategy, at least from the perspective of minimising problems during test delivery, is the offline model. In this case, the test delivery application and associated content are installed and run locally on each computer used for testing. While this typically results in the lowest failure rate, it presents the logistical burden of installing the testing software and collecting the results of the test upon completion. This difficulty can be reduced by running the testing software from a removable device, such as a USB memory stick, and configuring the software in such a way that no installation step is required. For a number of previous assessment projects, including PISA 2012, we have followed this offline approach, using a technology called "Portable Apps" (http://portableapps.com/). With this setup, we install all the software needed to run the test, including the web browser, web server and database server, on the memory stick and run the test directly from that device. The test content is stored on the memory stick, as are the results of the test session.

29. We propose to also offer a variation on this model for countries with more robust Internet connectivity in their schools. In this case, the testing software could be downloaded to the school computers from the Internet (either from the PISA Portal or from a server in the national centre), and copied to the computer hard drives. This would be a copy operation only and would not require any setup to be performed that might require increased access privileges (e.g., "administrator rights"). When the test is completed, the testing software would be configured to upload test results directly to the PISA Portal. In this case, there would be no need for test proctors to visit each computer to collect the test results or to go through the process of uploading those results from the memory stick to the PISA Portal. This will greatly streamline activities within each school, as well as reduce the likelihood of memory sticks (and consequently test results) getting lost in transit back to the national centres. The memory sticks would be available as a fallback, in case Internet connectivity is not available on the day of the test.

**Processing the PISA results**

30. The ability to automatically score student responses is a key potential benefit of moving PISA to a more fully computer-based platform. Automatic scoring increases the reliability of scores, both within and across countries; reduces country burden in terms of having to recruit, train and supervise scorers; and, most importantly, enables various forms of adaptive testing, which may improve measurement by providing more targeted assessments of student skills. Data coming from computer-scored items is cleaner, with fewer errors and inconsistencies that would be found in a paper-based test. Finally, the data can be tracked automatically to determine if parts are missing or corrupted.

31. The TAO platform is presently capable of automatically scoring a range of item types that fall into two categories: selection and entry items. These items are within the standard item types discussed previously. Therefore, when test developers author these items they will be able to specify the correct answer(s) and scoring occurs automatically.

32. The ability to score selection items means that several types of PISA trend items can be automatically scored. Clearly, those items that already in a multiple-choice format can be easily transitioned to automatic scoring. For items in which the correct response is located in the stimulus, a number of options exist that would allow computer scoring. These include asking students to

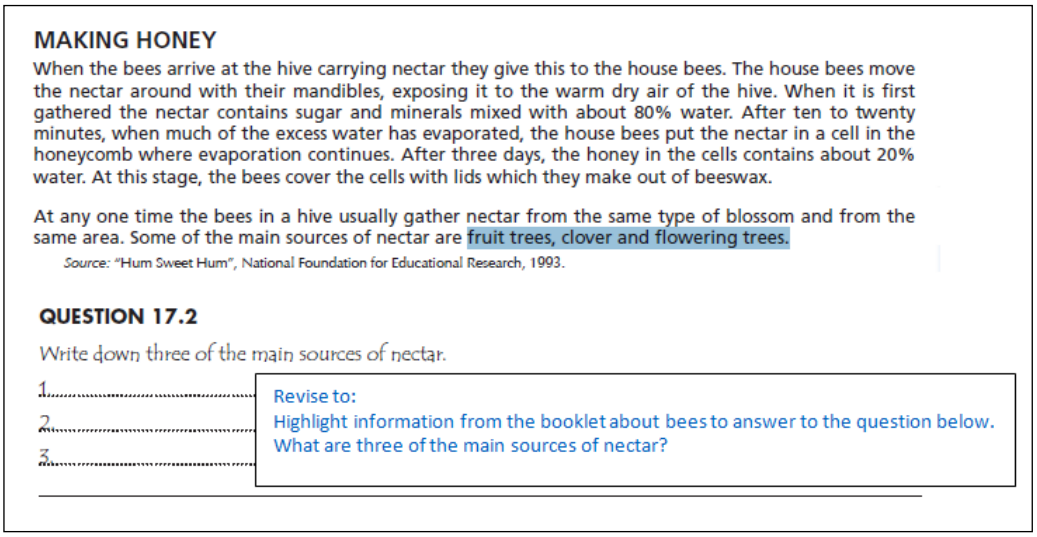
highlight their response in a passage or click on an image, a cell in table, or on a location on a graph in order to indicate their response.

33. For PISA[1] this means that the Core 1 and Core 3 contractors could consider click or highlight response modes for items including:

- Open-ended reading items that don't require a judgment - such as those where students have to access and retrieve information or, in some cases, integrate and interpret information.

  *Final Section of "Bees" Passage and Question As It Could Be Revised for Highlighting Response*

  *("Take the Test" page 57)*



**MAKING HONEY**

When the bees arrive at the hive carrying nectar they give this to the house bees. The house bees move the nectar around with their mandibles, exposing it to the warm dry air of the hive. When it is first gathered the nectar contains sugar and minerals mixed with about 80% water. After ten to twenty minutes, when much of the excess water has evaporated, the house bees put the nectar in a cell in the honeycomb where evaporation continues. After three days, the honey in the cells contains about 20% water. At this stage, the bees cover the cells with lids which they make out of beeswax.

At any one time the bees in a hive usually gather nectar from the same type of blossom and from the same area. Some of the main sources of nectar are fruit trees, clover and flowering trees.

*Source: "Hum Sweet Hum", National Foundation for Educational Research, 1993.*

**QUESTION 17.2**

Write down three of the main sources of nectar.

1.....

2.....

3.....

Revise to:
Highlight information from the booklet about bees to answer to the question below.
What are three of the main sources of nectar?

- Some items representing more complex information processing skills such as the example below based on a text about five cloned calves; this item was classified as a Drawing/Evaluating Conclusions item.
  "*Calf Clones Test*" (http://pisa-sq.acer.edu.au/showQuestion.php?testId=2300&questionId=9)
  Existing paper and pencil format

Which of the following statements is/are true? Enter Yes or No for each.

| Statement | |
|---|---|
| All five calves have the same type of genes. | Yes/No |
| All five cales have the same sex. | Yes/No |
| The hair of all five calves has the same colour. | Yes/No |

Computer-based variation that could be automatically scored.

Which of the following statements is/are true? Click on Yes or No for each.

| Statement | | |
|---|---|---|
| All five calves have the same type of genes | ○ Yes | ○ No |
| All five calves have the same sex. | ○ Yes | ○ No |
| The hair of all five calves has the same colour. | ○ Yes | ○ No |

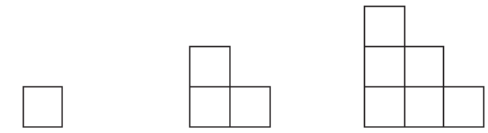
[1] Sample PISA items taken from "Take the TEST: Sample Questions from OECD's PISA Assessments" OECD, 2009.

34. Entry items are those for which students must provide a response. For these items there are no provided choices or available information in the stimulus that can be identified (highlighted, clicked, etc.) to indicate a response. Students must write, or in a computer-based test, type a response. The entry items that can be most easily computer scored are those requiring a numeric response as in the two examples below. Thus, many of the mathematics items can be converted to items that can be scored by the TAO platform. Test developers will need to define specific correct responses or a range of correct responses in order for the computer to score a given student answer.

- Items requiring a numerical response
  *Mathematics Item* (*Take* the Test" page 124)

**QUESTION 26.1**

Robert builds a step pattern using squares. Here are the stages he follows.




As you can see, he uses one square for Stage 1, three squares for Stage 2 and six for Stage 3.

How many squares should he use for the fourth stage?

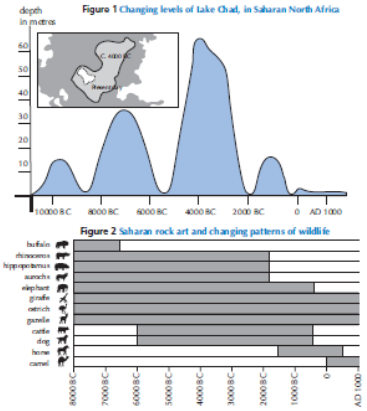
Answer: .....squares.

*Reading Item Requiring Numerical Response ("Take the Test" page 17)*

Figure 1 shows changing levels of Lake Chad, in Saharan North Africa. Lake Chad disappeared completely in about 20,000 BC, during the last Ice Age. In about 11,000 BC it reappeared. Today, its level is about the same as it was in AD 1000.

Figure 2 shows Saharan rock art (ancient drawings or paintings found on the walls of caves) and changing patterns of wildlife

*Source*: Past Worlds: The Times Atlas of Archaeology, Times Books Limited 1988




Use the above information about Lake Chad to answer the questions below.

**QUESTION 1.2**

In about which year does the graph in Figure 1 start?

.....

35. When entry items require a text response, automatic scoring - particularly in an assessment conducted in multiple languages – becomes problematic. From a technical perspective, there is no problem. A text-based response can be easily scored, but only if test developers can specifically identify a finite list of correct responses. The issue becomes one of accounting for variations in wording, spelling errors and typographical errors – all things that human scorers can easily evaluate but which computers cannot. If the response is limited to a word or two, countries could be asked to

provide a list of correct variations of the response which would make these possible to computer score. But any variations from the specified list would be scored as incorrect by the computer. If the capability to score single-word responses was deemed important by the framework and test developers, we would work in concert with the psychometric experts to determine if reliable scores could be obtained for this type of entry item. One possibility would be to use the results of the field test to create the list of correct responses that would be considered for the main study.

36. Entry items requiring a more elaborated response, such as providing an explanation, justifying a selection, or "showing your work" for a mathematics problem, obviously present the greatest challenge. The platform will provide a means to export the response data for items that cannot be machine scored so that is can be provided to scorers (at a minimum, via printed responses) and a system for entering the scores into the TAO database.

37. For more complex items, the TAO platform supports tracing common user actions (such as clicks, mouse movements, text entry, etc.) and logging those via an API. Working with programmers, item developers can also define their own events that will be traced just as common events would. These custom events can be related to item content and will allow us to capture specific actions or processes of interest. To assist in processing these logs, the platform includes the TAO Event based Scoring using Log Analysis service, or TESLA. TESLA is based on four elements:

- An event log model
- A variable description language to define the components of the score and express the rules for processing the results
- The TAO Event Processing Engine, which is a set of services for analysis of the logs, including a complex pattern matching service
- A visual interface that simplifies the definition of variables