

José-Guillermo Hernández-Calderón, Ph.D.
guillermohernandez02@uv.mx

July 5th, 2024

The elephant in the room...

- Maybe my English is not very good looking
- I'm only been related to language assessment for five years now
- I'm very nervous



Today's talk

- Computer Assisted Language Testing (CALT)
- A little bit of theory and history
- The Computer-Assisted part of CALT
- Lessons learned over five years

CALT

Computer-Assisted Language Testing

“use of computer devices for language assessment”

CALT

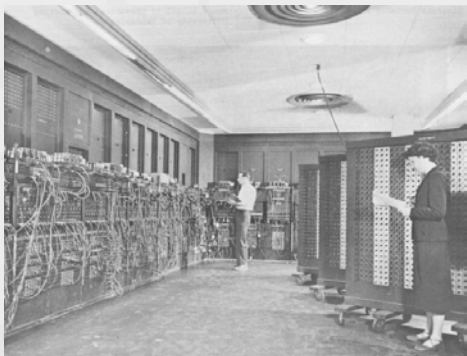
Computer-Assisted Language Testing

- dates back to **1935** when the IBM model 805 was used to **score** objective tests in the United States of America.
- It helps reduce the labor-intensive and costly business of scoring thousands of tests each year.



Evolution of computers in testing

- The first computer filled a 20 by 40 foot room and weighed 30 tons.
- Now we carry with computers in our pockets everyday.



Part One Numbers 1 – 4

INSTRUCTIONS: For each number, fill in the box to match each statement (1 – 4) with the correct notice (A – F). There are TWO extra notices.

EXAMPLE

6. This notice advertises a new, better workout routine for young adults.

CORRECT ANSWER

G. "Get go!" exercise activities help you feel better and breathe properly.

1. The best vacation homes are frequently available for people to buy here.

2. This tasty snack might seem sweet but has some natural ingredients.

3. You can find places to rent that are suitable for families with kids.

4. These items have less of some things but might make people healthier.

A BEST VACATION HOMES OFFERS A REMARKABLE SAFE LOCATION THROUGHOUT THE YEAR.

B LOW-WEIGHT SNACKS WITH NATURAL INGREDIENTS – 80% KIDS APPROVE!

C FEEL HEALTHIER! NATURAL SNACKS IN PORT CONVENIENT NEW! CUSTOMER DISCOUNTS!

D PROPERTY MANAGER SEARCH! ATTRACTIVE! MODERN! AFFORDABLE! TO PURCHASE OR RENT! FURNISHED! YOURS TO KEEP!

E TRY OUR BOLD TASTE! REMARKABLE! 90% KIDS LOVE! BEST SNACKS TO LOVE! (SEE US AT THE STORE!)

F LOOKING FOR HEALTHY ACCOMMODATION? MODERN! AFFORDABLE! TO PURCHASE OR RENT! FURNISHED! YOURS TO KEEP!

- Computers were used during the scoring process and now are used in all the evaluation process

Issues

- Security
- Technical expertise
- Inherent differences between paper-based tests and computerized tests
- Privacy concerns

Advantages

- Works perfectly with a small staff
- It saves time and money
- Immediate results and feedback
- Enhanced interaction/
Personalized experience
- Endless opportunities for improvement and innovation

Pathan, M. M. (2012). Computer Assisted Language Testing [CALT]: advantages, implications and limitations. Research Vistas, 1(4), 30-45.

CALT

Computer-Assisted Language Testing

Computer-Assisted

Test delivery platform User experience
Usability Deployment Databases Software design
Scalability Validity Networks High availability
Computer-Assisted
Digital literacy System security User interaction
User-Centered Design Data storage
Device compatibility Proctoring software

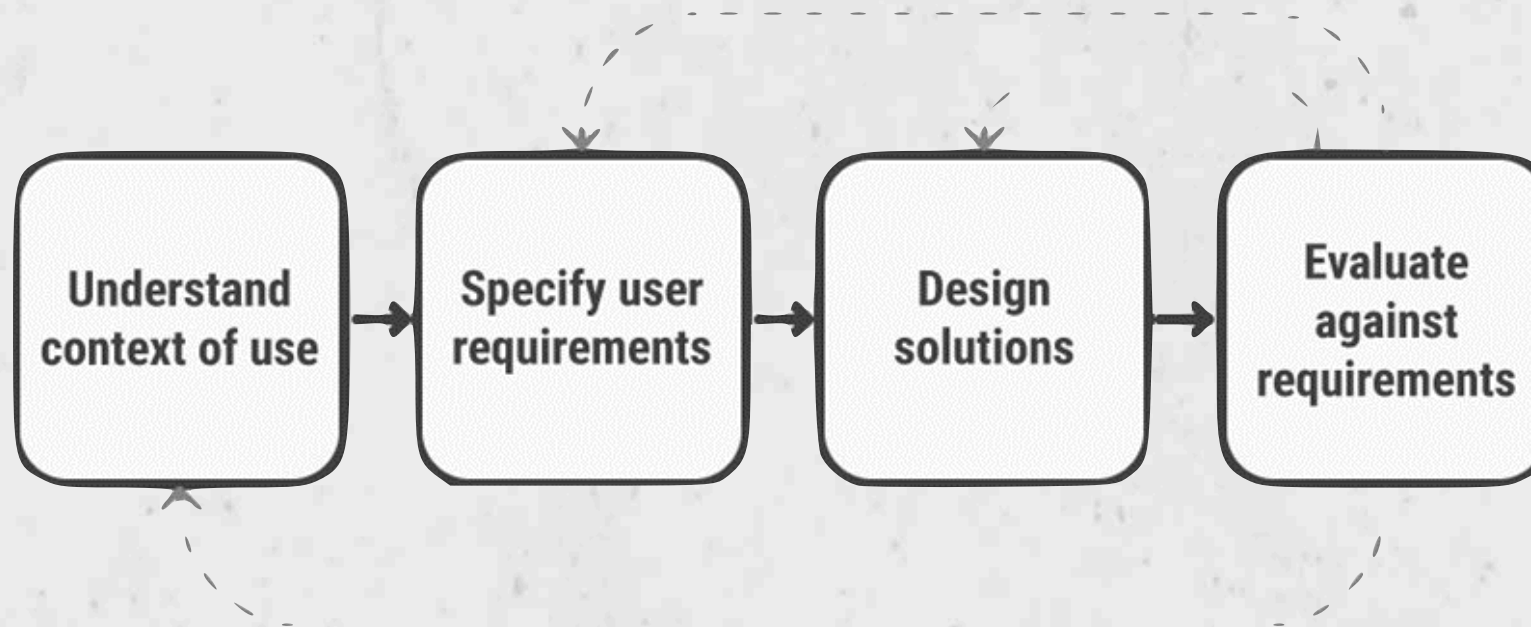
Software design



Process of conceptualizing **how a software system will work** before it is implemented...

Design methodologies help developers plan, produce, and test software

User-Centered Design (UCD)

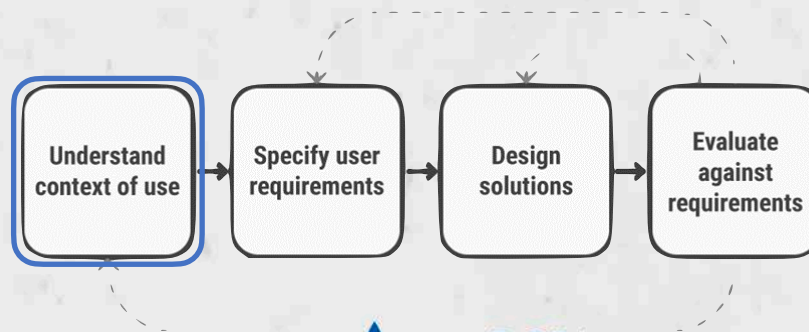


iterative methodology in which
designers look at users and their needs
in each phase of the design process

Understand context of use

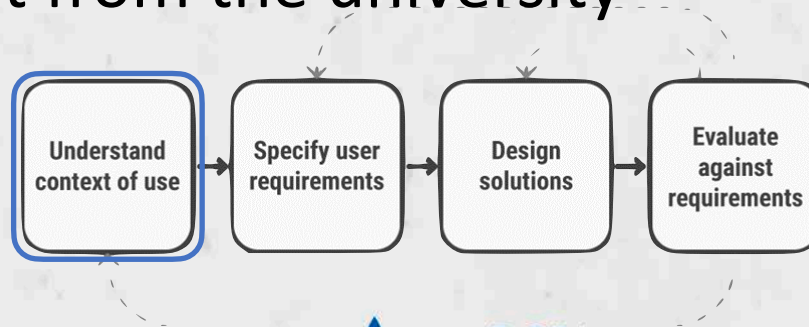
“What people say, and what people do, and what people say they do, are entirely different things.”

Margaret Mead, Ph.D., Anthropologist

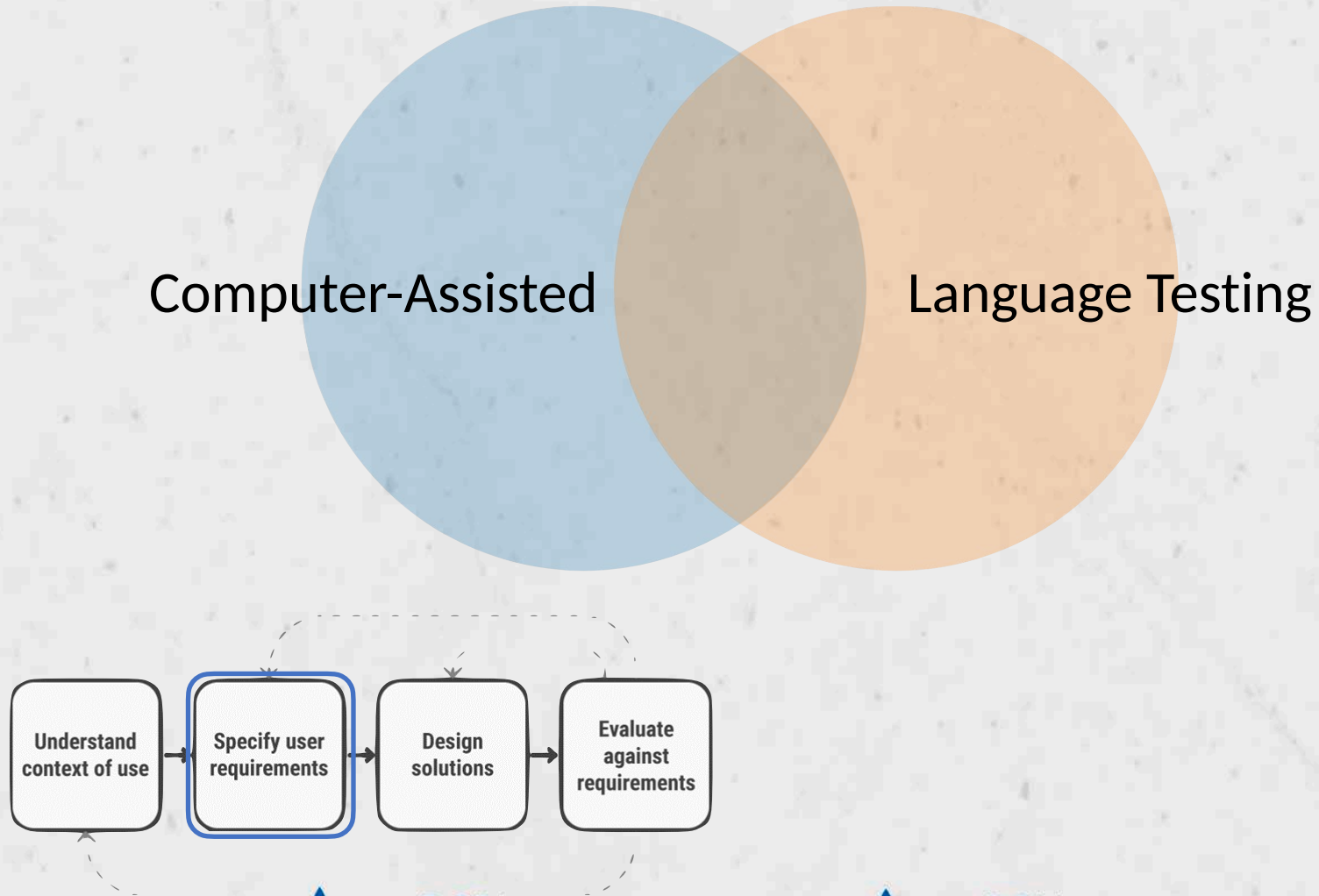


Understand context of use

- EXAVER evaluates 3 different levels A2, B1 and B2 in three exams: EXAVER 1, EXAVER 2 and EXAVER 3
- **+4000** exams given **per year**
- **Standardized** tests
- Candidates come from various contexts, not just from the university



Specify requirements

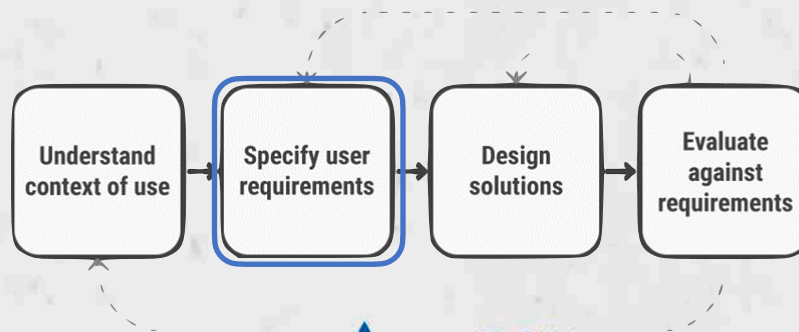


Specify user requirements

Some related with language testing process,

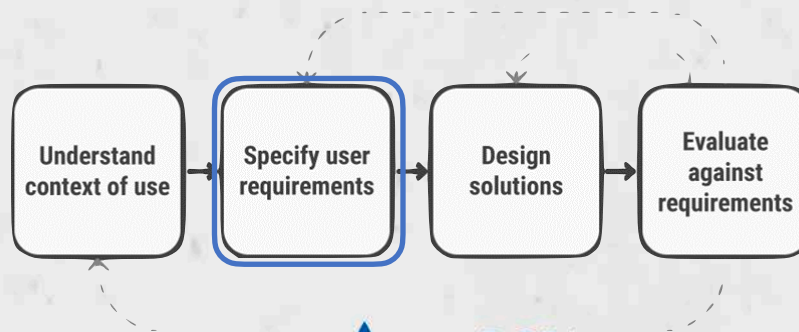
- Test design
- Test specifications
- Validity
- Reliability

Those related with the computer implementation

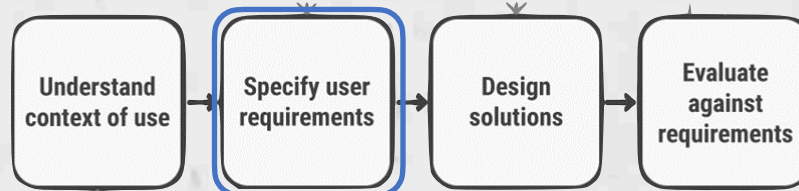
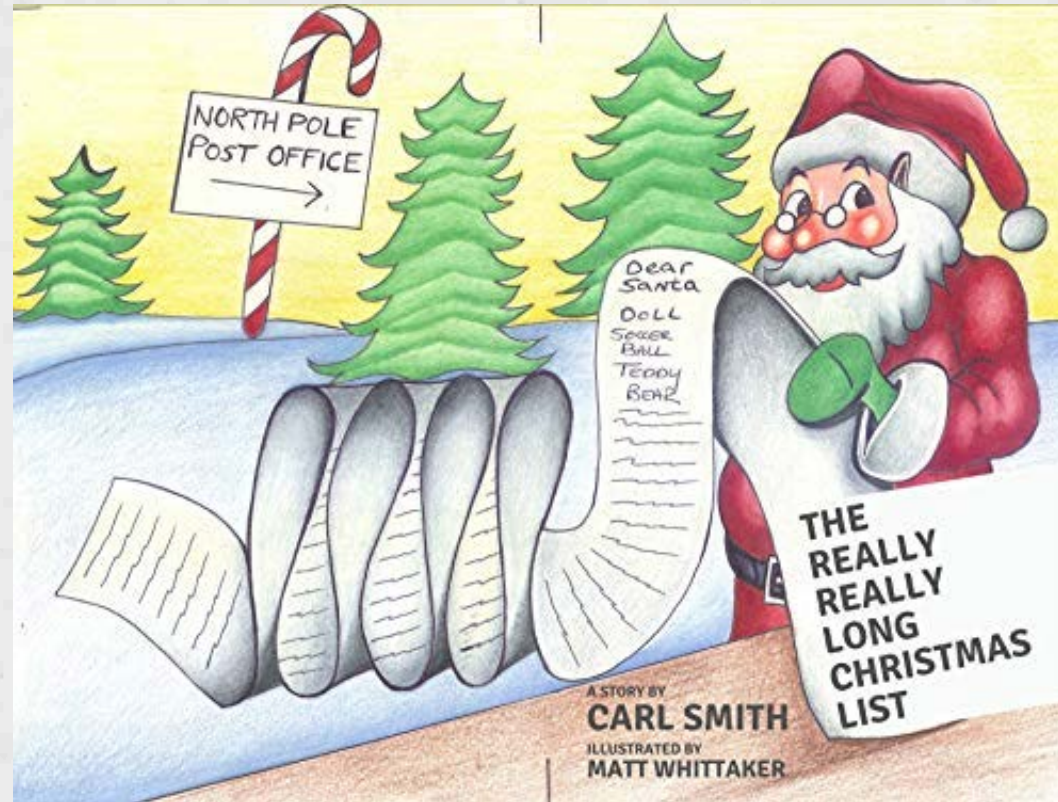


Specify user requirements

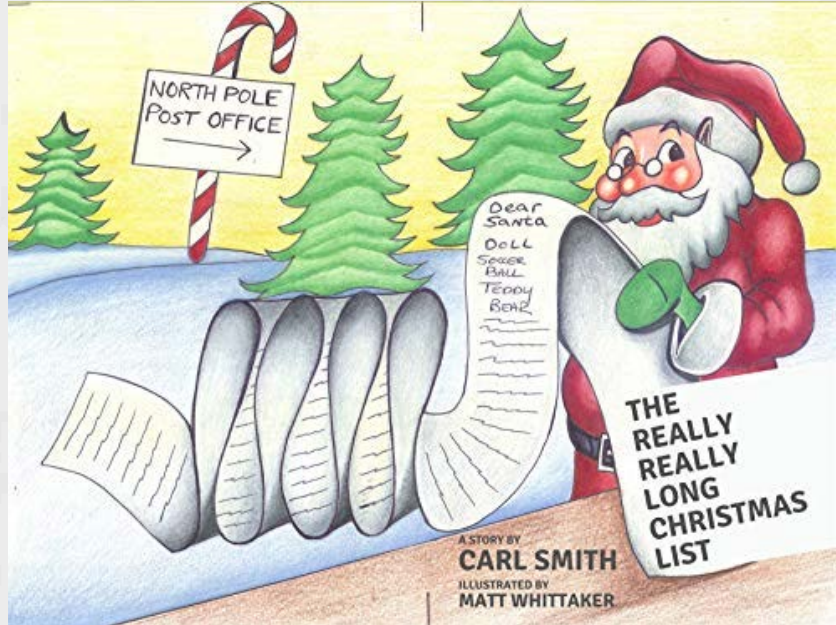
- Control test-takers access
- Security of items and data
- Improve evaluation process
- Automatic scoring
- Consistent design
- Test-taker identification
- System and hardware requirements
- Administration of exam rules
- Candidates digital abilities
- Storage candidates' answers and actions



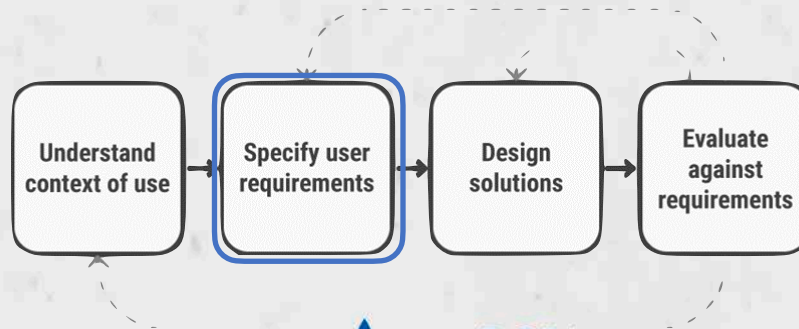
Specify user requirements



Specify user requirements



At the end, majority of concerns are in the table but **a balance** between requirements and what is it possible to do considering time and other resources like staff, money, etc... **is required to implement effective evaluation tools**



Design solutions

From paper-based to high-fidelity prototypes



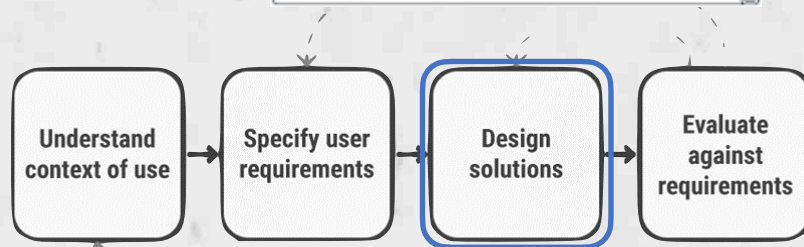
Registro de calificación examen oral

Número de candidato: 69000
Nombre del alumno: Para pruebas

Examiner only

	1	2	3	4
Pronunciation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Vocabulary Range	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Grammatical Accuracy	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Effectiveness in Communication	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

☐ **No presentó**



Improve evaluation process

Design solutions

Registro de calificación examen oral

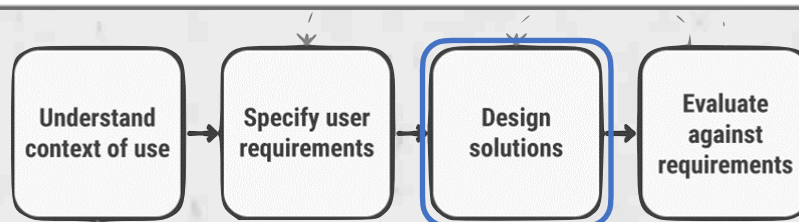
Número de candidato: 69000

Nombre del alumno: Para pruebas

Examiner only

	1	2	3	4
Pronunciation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Vocabulary Range	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Grammatical Accuracy	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Effectiveness in Communication	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

☐ **No presentó**



EXAVER 3

octubre 2019

[Descargar resultado](#)

C

Acreditado

INTERPRETACIÓN DEL RESULTADO:

Puede leer textos sobre temas relacionados con su campo e interés con un nivel de comprensión satisfactorio. Puede transmitir información e ideas sobre temas abstractos y concretos, verificar información y preguntar o explicar problemas con una precisión satisfactoria. Puede comunicarse con cierta confianza en asuntos familiares rutinarios y no rutinarios relacionados con sus intereses y campo profesional. Puede intercambiar, verificar y confirmar información. A menudo es incapaz de entender y producir formas gramaticales intermedias-altas ni iniciar o tomar turnos adecuadamente.

DESGLOSE POR HABILIDADES:

PAPER ONE		PAPER TWO	PAPER THREE
Reading	Writing	Listening	Speaking
<div></div>	<div></div>	<div></div>	<div>Speaking 20%</div>

Consideración:
Los resultados publicados en esta página web son únicamente de carácter informativo y no tienen validez oficial.

[Ver detalles de inscripción](#)
[Constancia de acreditación](#)

Design solutions

Test-taker identification



EXAVer Online - Exámenes de Certificación de Lengua Inglesa

Exámenes de Certificación de Lengua Inglesa

Acceso al examen escrito

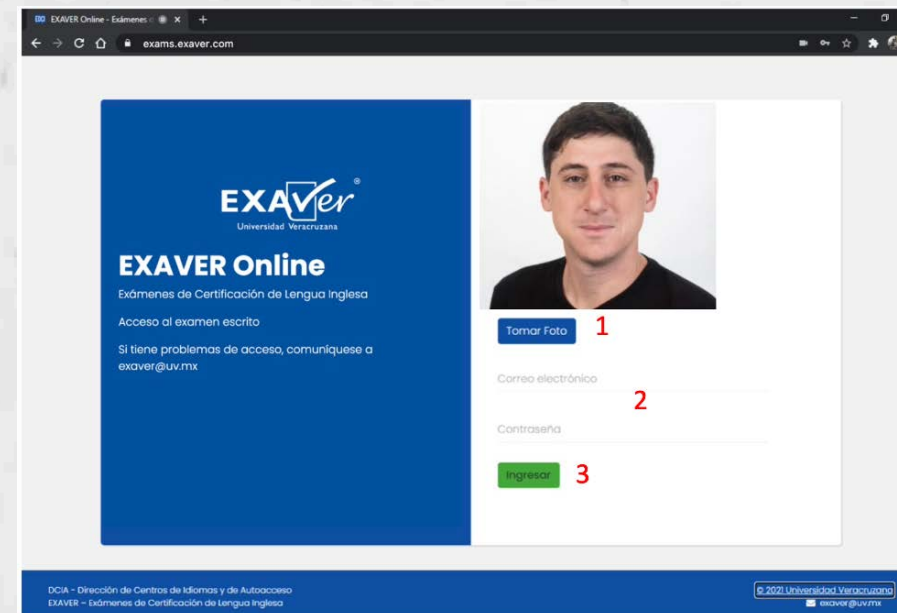
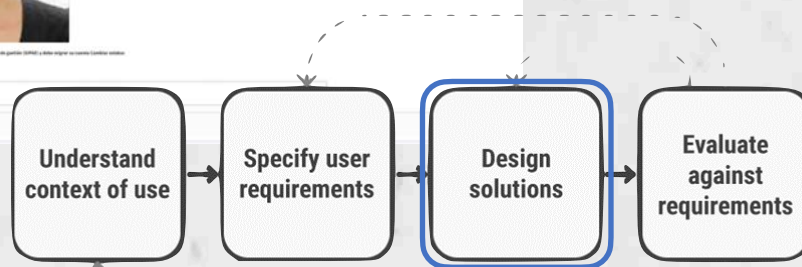
Si tiene problemas de acceso, comuníquese a exaver@uv.mx

Tomar Foto 1

Correo electrónico 2

Contraseña

Ingresar 3



Assessment process was modified to guarantee test validity and security of the administration process.

Design solutions

Part One **Numbers 1 – 4**

INSTRUCTIONS: For each number, fill in the box to match each statement (1 – 4) with the correct notice (A – F). There are TWO extra notices.

EXAMPLE

5. This notice advertises a new, better workout routine for young adults.

CORRECT ANSWER

G

1. The best vacation homes are frequently available for people to buy here.

2. This tasty snack might seem sweet but has some natural ingredients.

3. You can find places to rent that are suitable for families with kids.

4. These items have less of some things but might make people healthier.

A

B

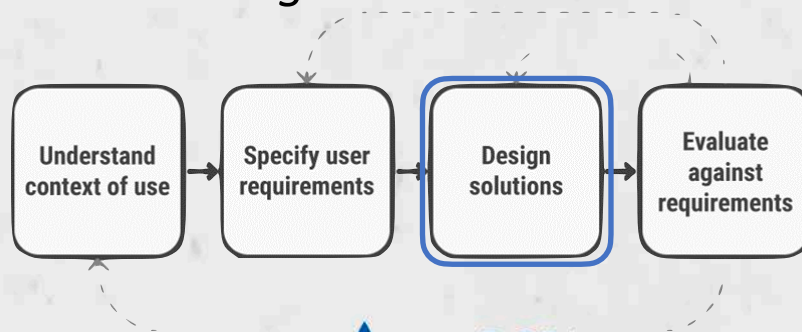
C

D

E

F

Consistent design

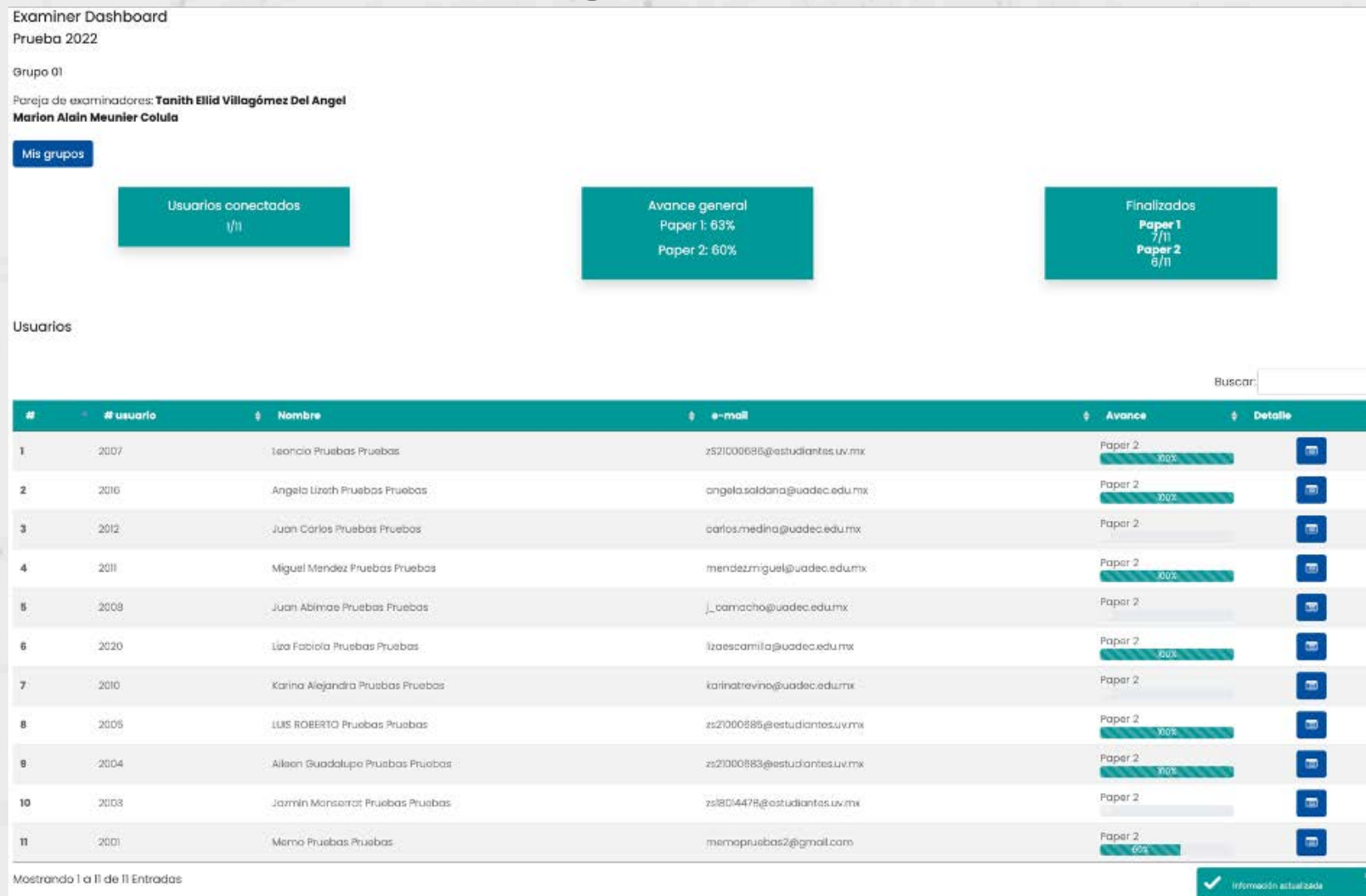


Behavior		Definition	Event	
Class	Subclass		Name	Reason
Normal		Consistent with what the test-taker should do during the test	Access allowed	Test-takers need to login to the system
			Access granted	Test-takers need to access to a specific test
Abnormal	Low	Can happen in an acceptable number of occurrences	Active session	A login attempt was made to a logged in user.
	Medium	Can happen in an acceptable number of occurrences	Exit browser	Test-taker stopped interacting with the Google Chrome® browser
			Returns to browser	The user started interacting with the Google Chrome® browser again.
	High	Undesirable, Test-takers are not expected to behave this way because they give clues of suspicious activities	Tab change	Only the official test web page should be opened in one browser tab
			New tab created	No additional tabs are needed for the test
			URL change	Only the official test web page should be opened

Storage candidates' answers and actions

User actions data logs

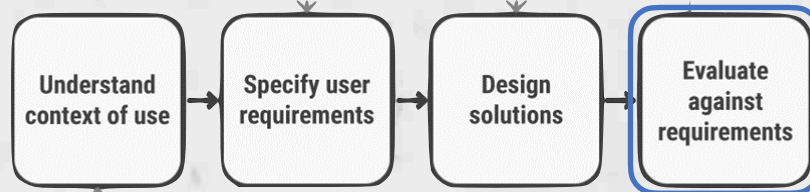
Data Visualization Tools



Were computerized test requirements fulfilled?

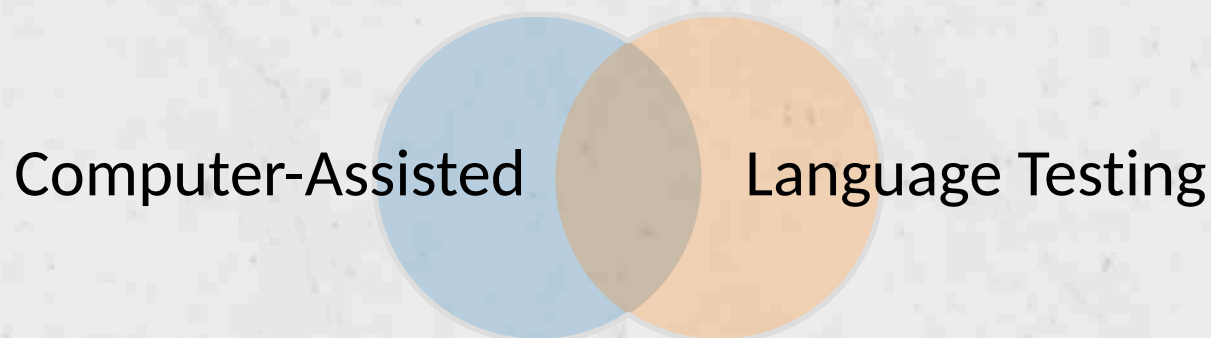
- Evaluation sessions are held with users and experts
- Corroborate that the software works as conceptualized.

Process of conceptualizing
**how a software system
will work** before it is
implemented...



Final remarks

- The **design** of CALT tools is **highly influenced by context**
- **Technology and computers in CALT aren't the core**; they serve as a tool that supports language assessment processes



- **Balance is important**, but don't forget to ...

Current trends

AI-powered tools

Write & Improve
with Cambridge



Very good. Your new level is B2. Your writing is improving. Use the feedback and think about ways to improve more. Make changes. Click Check again when you are ready.

Your CEFR level for this question is at the top of 'Task Help'

Task help

History

Help

Level B2

Images

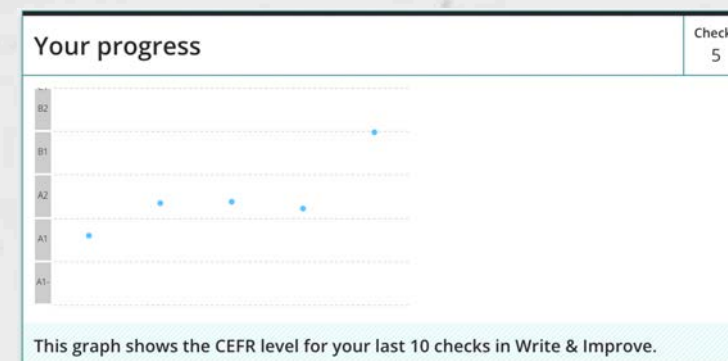
W&I feedback

Changes

You raised your level to B1. Very good! Your writing is improving. Use the feedback and think about ways to improve more. Make changes. Click Check again when you are ready. Or, return to Workbooks and start a new task.

Your CEFR level for this question is at the top of 'Task Help'

One of the most important modern functions of zoos is supporting international breeding programmes, particularly for endangered species. In the wild, some of the rarest species have difficulty in finding ☆ mates and breeding, and they might also be threatened by poachers, loss of their habitat and predators. A good zoo will enable these species to live and breed in a secure environment. In addition, as numbers of some wild species drop, there is an increased danger of populations becoming too genetically similar. Breeding programmes provide a safeguard: zoo-bred animals can be released into the wild to increase genetic diversity.



Current trends

AI-powered evaluations

Linguaskill ▶▶

Hallo 

 SmallTalk2Me

Current trends

Automatic Item Generation



Everybody asks about the item generation process—here it is.



Step 1

The content used for item generation is identified by subject-matter experts



Step 2

An item model, which is like a template, is created to specify where the content is placed in the generated item



Step 3

Computer algorithms place the content from step 1 into the item models in step 2

Current trends

Computer-Adaptive testing

Computer-adaptive testing of listening and reading skills

Computer-adaptive testing

In an adaptive test, each question is selected by the computer based on a candidate's performance in the previous questions. In general:

- ✓ If a candidate gets a question right, the computer will choose a harder question.
- ✗ If a candidate gets a question wrong, the computer will select an easier question.

To summarize

- Use of Artificial intelligence and data analytics
- Customized assessment tools
- AI-powered evaluations
- Adaptive testing methodologies
- Natural Language Processing

José-Guillermo Hernández-Calderón, Ph.D.
guillermohernandez02@uv.mx

July 5th, 2024