




Social Integration exams in the Netherlands

Marli Tijssen

CINOP

Overview:

- Government policy
 - Social Integration exams
 - Test construct, test technology
 - Automated scoring
 - Questions
- 
- Three white, semi-transparent diagonal bars are positioned in the bottom right corner of the slide, pointing upwards and to the right.

Government policy

- ...- 1995 no policy in relation to integration or citizenship
- 1998 Law on integration of Newcomers:
 - Responsibility of the Dutch society
 - Obligatory lessons of Dutch and knowledge of Dutch society
 - No exams but certificates.
- 2006 Law on integration:
 - Own responsibility of the migrant
 - Social Integration exam in home country
 - Social Integration exam in the NL

New migrants – Established migrants

- Established migrants:
 - Social Integration exam in the Netherlands
- New migrants:
 - Social Integration exam in home country
 - Social Integration exam in the Netherlands

Social Integration exam in home country

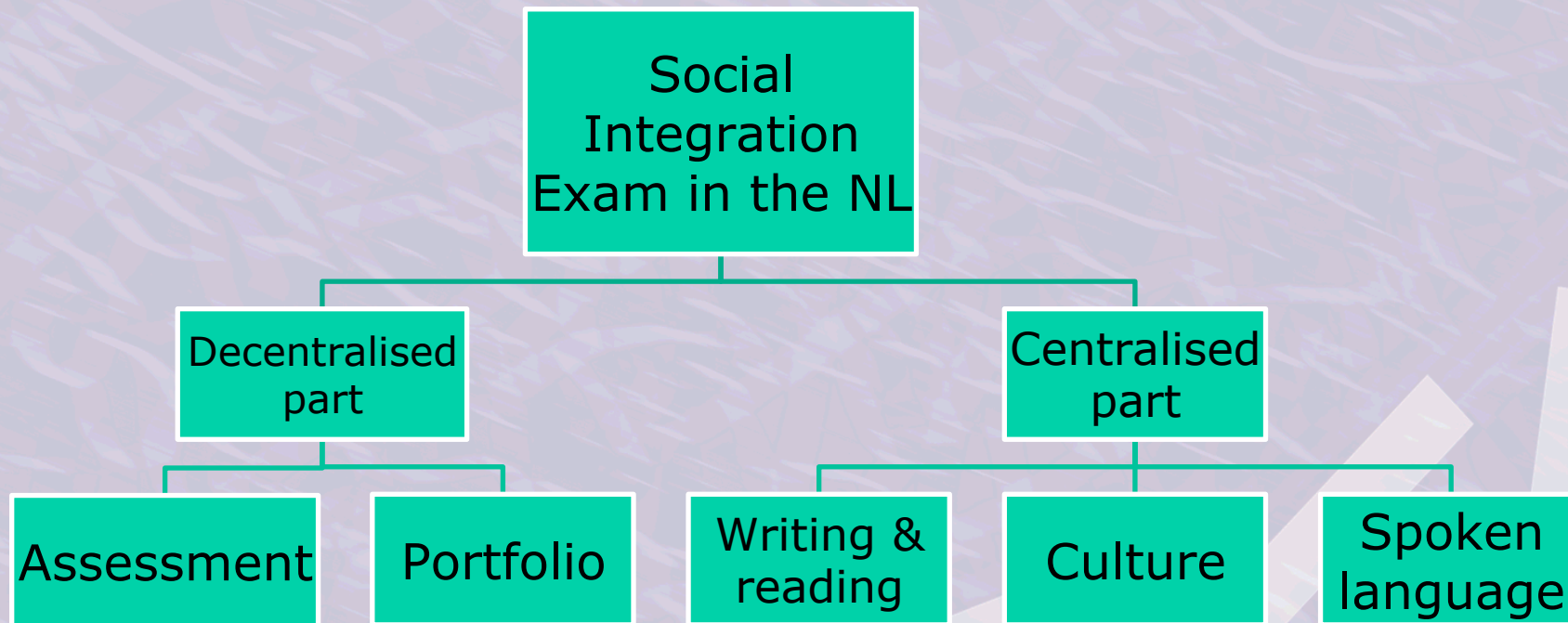
Social Integration
Exam
Dutch Embassy

Spoken
Dutch Test

Knowledge
Dutch
Society

Literacy and
Comprehensive
Reading

Social Integration exam for migrants already living in the Netherlands



Requirements for Language Test and Culture test

- Make a test of spoken Dutch and Knowledge of Dutch Society that:
 - Can be taken at any time, any place
 - Can be taken without written materials
 - Can be administered by embassy staff
 - Has immediate results
 - Is absolutely secure
 - Has results expressed on scale of Council of Europe (language test)
- And of course
 - Reliable, valid and all that jazz



Three Companies

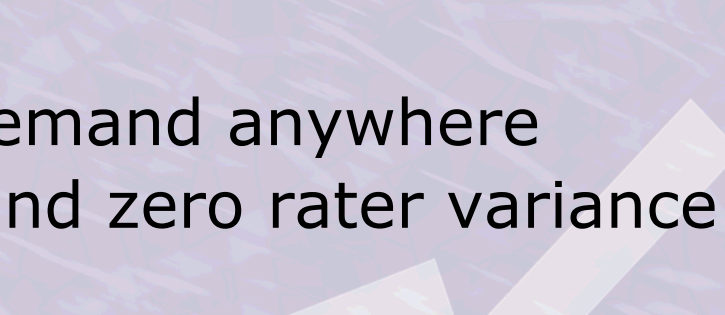


PKT's technical solutions

Tests that measures listening and speaking performance

- In direct interaction
- Using speech recognition & automated scoring

Advantages:

- Spoken, real time
 - Directly available on demand anywhere
 - No rater dependence and zero rater variance
 - Immediate results
- 
- A decorative graphic consisting of three overlapping, semi-transparent white rectangular bars arranged in a diagonal pattern from the bottom left towards the top right.

What is tested?

Facility in spoken language:

- Ability to track what is said
- Extract meaning in real time
- Formulate responses
- Produce responses at a conversational pace, that are relevant and intelligible

Item types TGN (N=50, scored 45)

Repeating

Listening

Sentence structure

Pronunciation

Fluency

Open question

Listening

Vocabulary (Receptive & productive)

Opposites

Listening

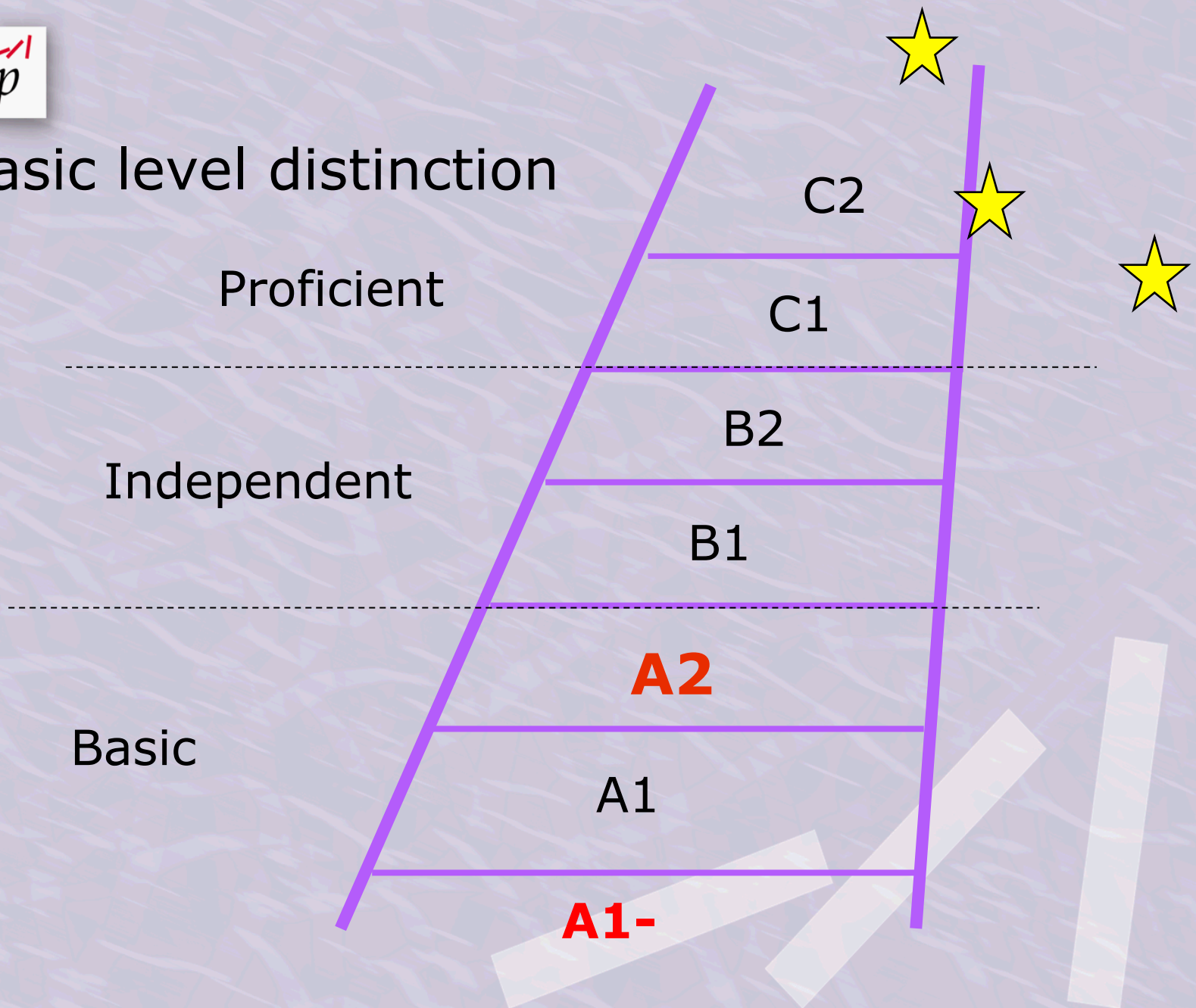
Vocabulary (Receptive & productive)

Story retelling

Only for validation



The basic level distinction (CEF)

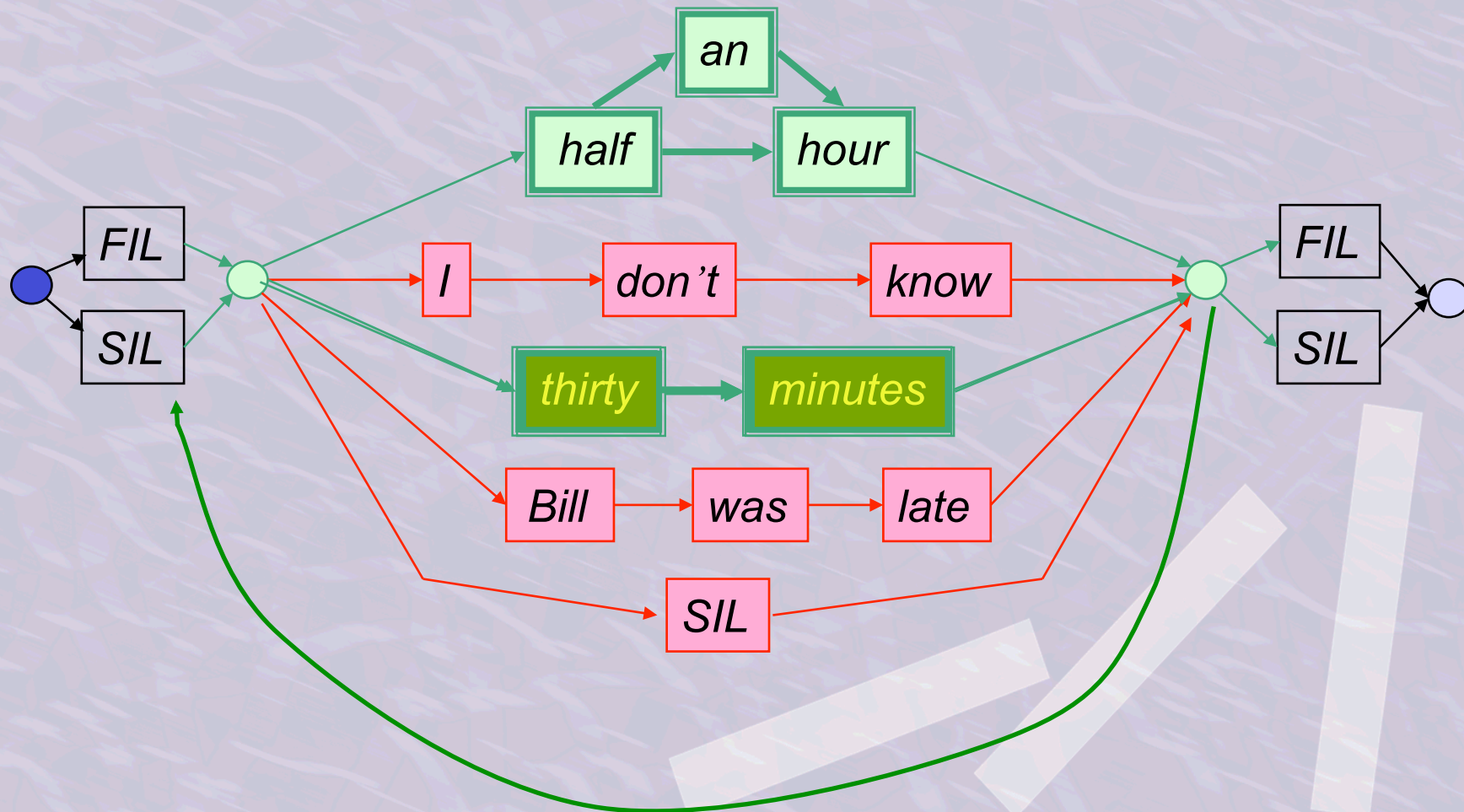


Example: open short answer question

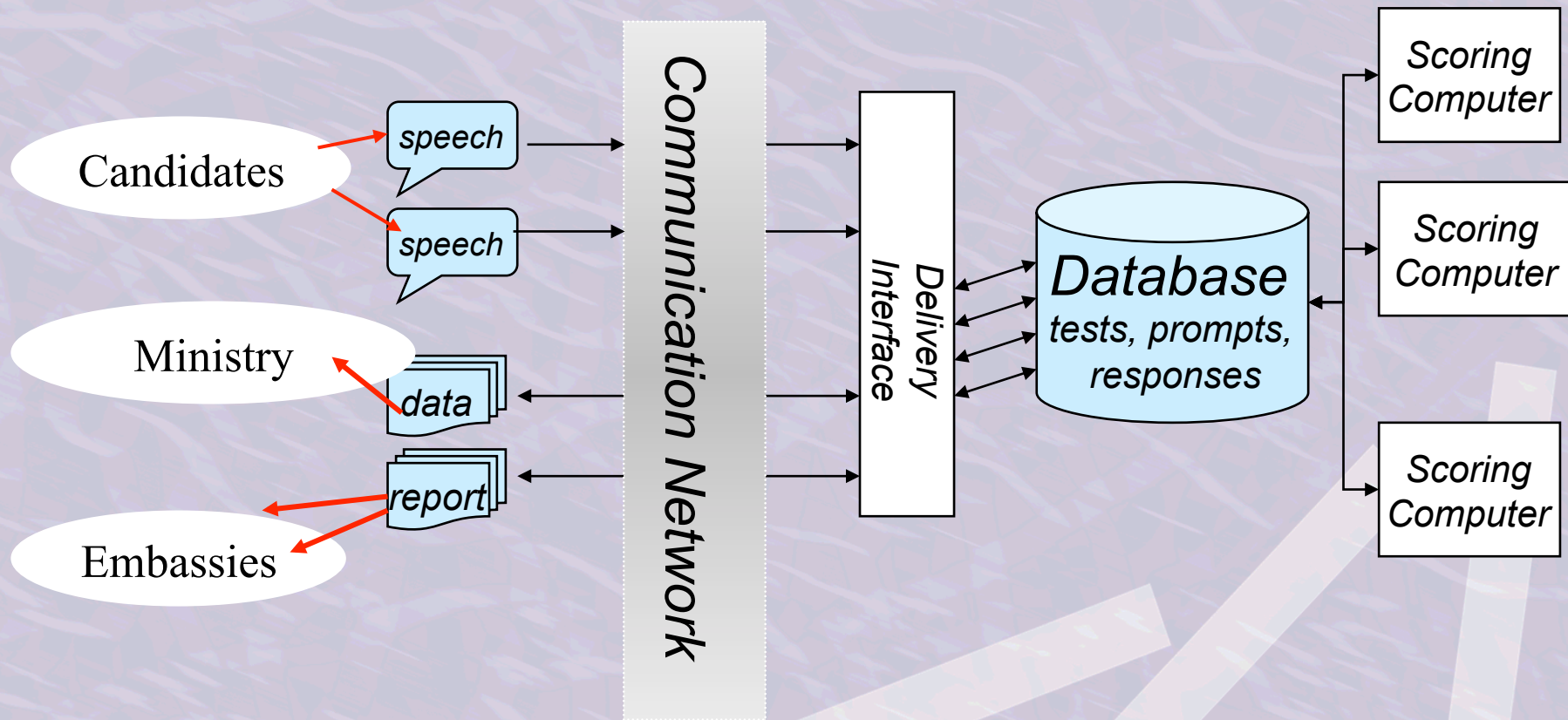
Bill arrived at 12:30 for a meeting at noon. How late was he?




Simplified Response Network



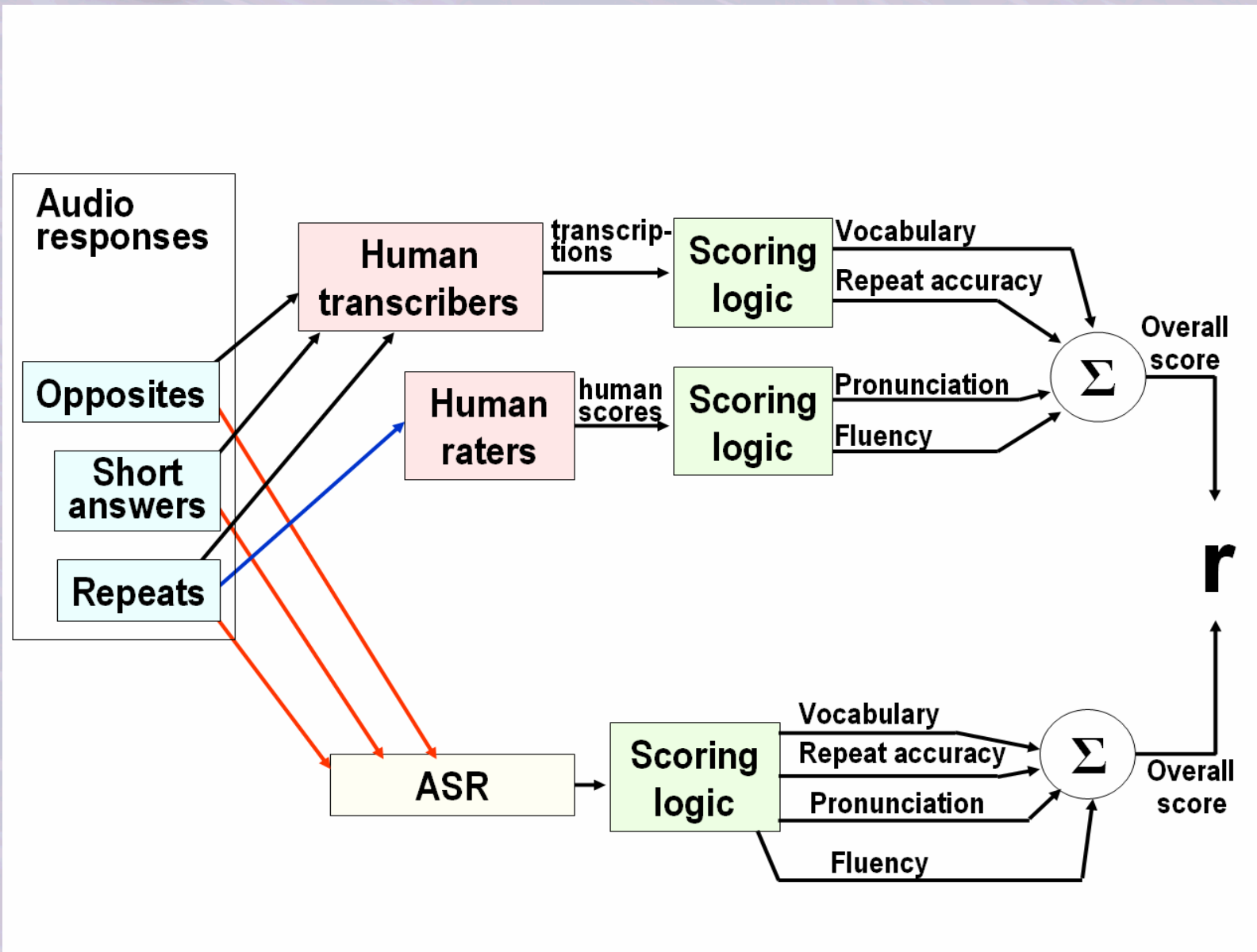
Testing Service Infrastructure



Automated scoring

- No human interference needed
 - Independent of time and place
 - Automated management reporting tool
 - Reliable and consistent
 - Efficient → reduced costs
 - Secure for loss and fraud
- 
- A decorative graphic consisting of three overlapping, semi-transparent white rectangular bars arranged in a diagonal pattern from the bottom left towards the top right.

Validation



Validation and reliability of machine scoring of the Language and Culture test

- $r = 0.93$ between human scores and machine scores of the Language Test
- $r = 0.94$ between human scores and machine scores of the Culture test

Political - and other impacts

- Criticism; not because of quality but because of political disagreement on immigration law
- Scepticism because of the innovative aspects of language testing
- Unfamiliarity with the CEF scales
- Competition from academic speech recognition community
- Unfamiliarity with speech recognition and technology

Independent research (TNO) on validation and reliability of automated scoring

- No substantial differences between human scores and machine scores on the Language and Culture test
- Candidates with strong foreign accents are not disadvantaged

Human rejudgement of machine scoring

Human rejudgement procedure:

1. Monitoring machine scoring and human scoring for a certain period
2. When no scoring is possible for reasons of status 1 (too much noise, no scoring possible)
3. Suspicion of fraud

Actual situation

2010:

- Language test: > 30,000 high stakes tests every year
- Culture test: approx. 10,000
- In Progress: Literacy and comprehensive reading test
A1 – CEF scale

Questions ?????





Marli Tijssen

+31(0)6-51690010

mtijssen@cinop.nl