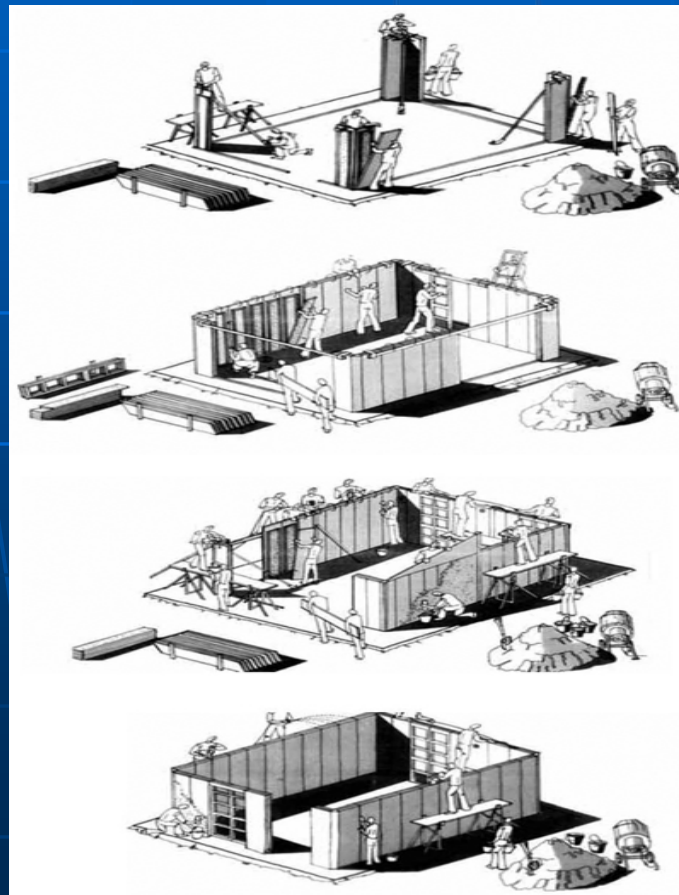


Che cosa dovete costruire ?



Test di lingua: definizioni

È uno **strumento di misurazione** avente come obiettivo:

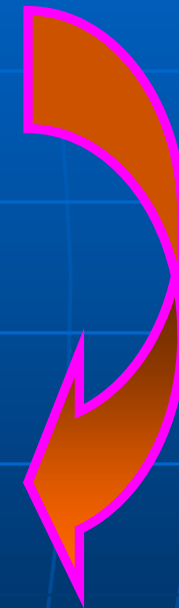
- elicitarne un campione di performance, un comportamento linguistico
- fornire cioè una misura che funga da indicatore della competenza

Tipologie di test

- di controllo (diagnostico)
- di piazzamento (di livello)
- di performance (task)
- a punti discreti (di solito grammaticale)
- integrativo (abilità integrate)
- normativo (per redigere per es. graduatoria)
- di competenza

Decreto e test A2

- Il riferimento è al QCER



pertanto

- il test A2 dovrebbe essere
test di competenza

Test di competenza: obiettivi

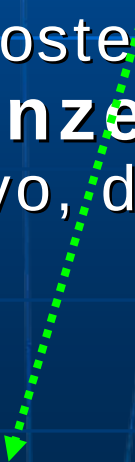
“elicitare, attraverso prove di verifica adeguate al livello ed appropriate all’utenza, campioni di performance che consentano di effettuare operazioni di inferenza sulla competenza linguistica, con l’obiettivo ultimo di prevedere la potenziale capacità d’uso della lingua in situazioni reali”.

(Specificazioni esami CELI)

Test di competenza: caratteristiche

- **Slegati da programmi didattici specifici**
- **Prove standardizzate** aventi duplice finalità:
 1. **sollecitare una performance**
 2. **assegnare alla stessa dei valori**
- Basati ed articolati sulla rigorosa descrizione degli obiettivi di apprendimento per i vari livelli del **QCER**
- Caratterizzati, sul piano della valutazione, da **almeno n. 4 misurazioni**

Alcune precisazioni terminologiche

- **Valutazione**: è un processo globale costituito dalla raccolta di informazioni attraverso procedure di **verifica** (relative sempre ai metodi) che portano alla sommatoria di più **misurazioni** (inerenti singole abilità e sotto abilità)
 - **Misurazione**: processo che porta a legare concetti astratti (**costrutto dell'abilità**) ad indicatori empirici (risposte candidati) attraverso compimento di **inferenze** e cioè attraverso un processo logico induttivo, da particolare a universale)
(Carmines e Zeller)
- 

... come, ad esempio ...

- **Modello pluridimensionale di Bachman**, costruito per gli esami CELI



Noto con l'acronimo **CLA**:
Communicative
Language
Ability

Misurazione e valutazione

- Non è l'apprendimento che si misura, ma i suoi risultati e cioè i comportamenti che lo manifestano

(Gattullo)

- (in questo senso) la misurazione precede la valutazione

(Visalberghi)

Peculiarità misurazioni L2 / LS

- Non solo inferenze ma anche → predizioni

Performance (particolare)

Inferenza



Abilità (generale)

Predizione

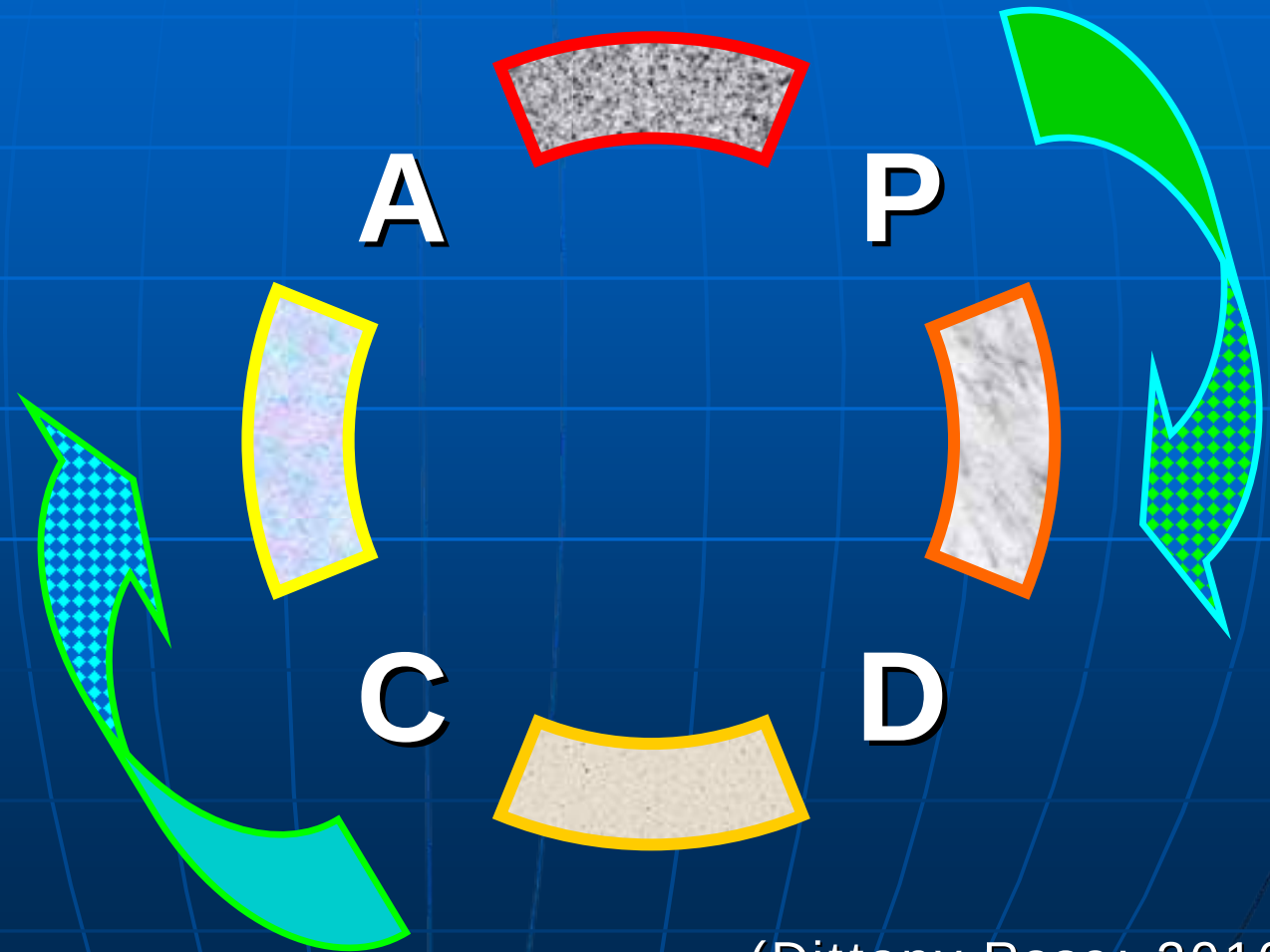


Competenza d'azione

“una valutazione che focalizza l'attenzione non su ciò che uno studente sa, ma su **ciò che sa fare con ciò che sa**”
G. Wiggins, definizione di **valutazione autentica o alternativa**

PDCA

- Plan
- Do
- Check
- Action



(Dittany Rose: 2010)

Sviluppo di un test: di cosa c'è bisogno

- Organigramma
- Documento ufficiale sulle politiche
- Specificazioni (due versioni)
- Linee guida per gli item writers
- Procedure di sicurezza
- Item bank
- Modulistica (iscrizione, registri, fogli firma ecc)
- Scale e criteri per l'assegnazione del punteggio
- Pacchetto per il training di amministrativi ed esaminatori
- Procedure di audit
- Software amministrativi
- Checklist
- Regolamento d'esame
- Software e hardware per la correzione / scansione
- Software per analisi statistica
- Procedure per la validazione, per il controllo della qualità, per la valutazione dell'impatto
- **Special arrangements**

Special arrangements ... per rimuovere le barriere

- Procedure per persone diversamente abili (1.7% iscritti CELI)
- 74 tipi di disabilità (cognitiva o fisica)
- Braille
- Extra time
- Accesso alle tecnologie
- Materiali per l'orale per daltonici
- Font e grandezza caratteri
- Formato A3



Ciclicità di un processo



Migranti:

- richiedenti Carta di lungo soggiorno
- non inseriti in contesto CTP (?)



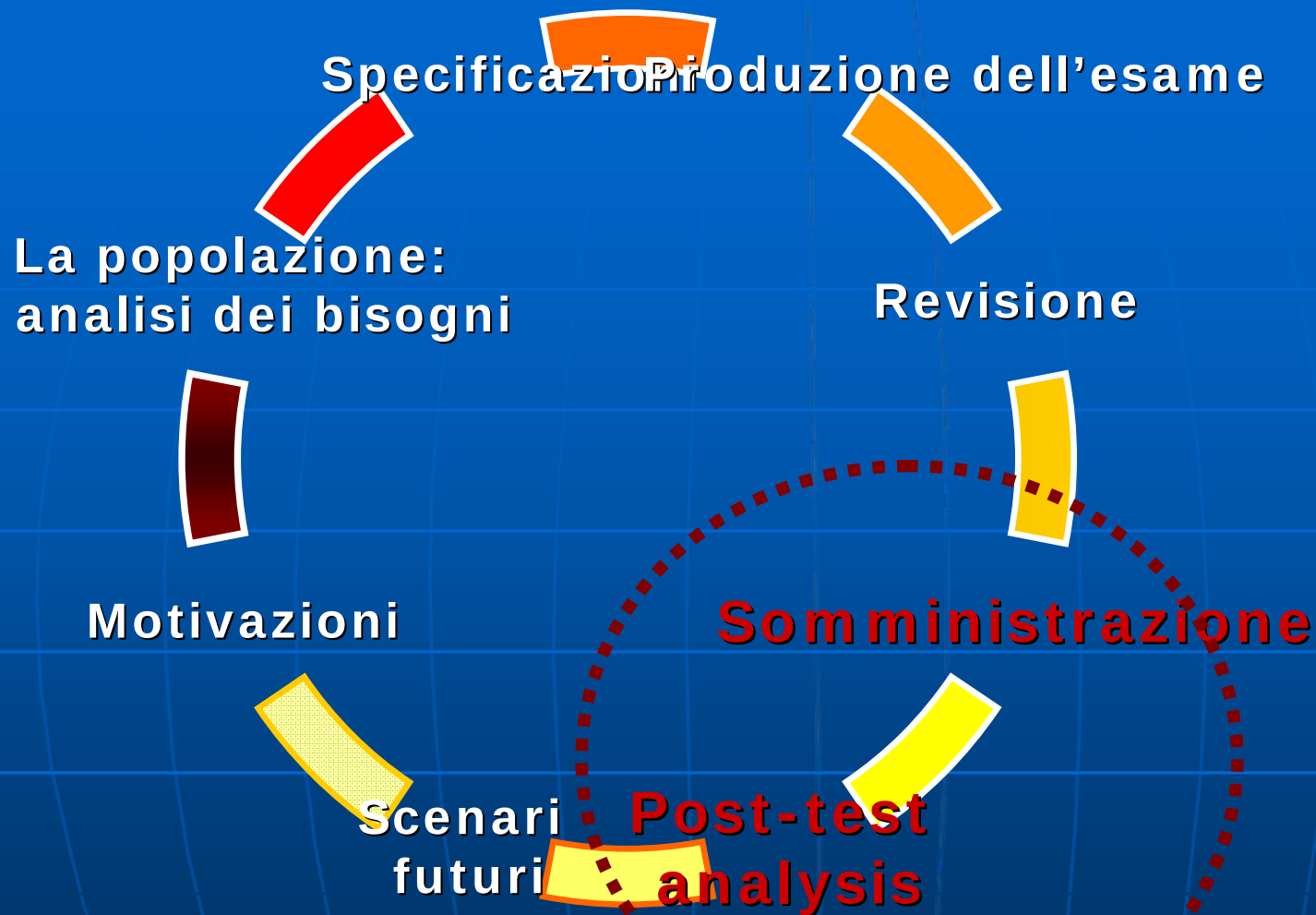
1. Sillabo ATS (?)
2. Vademecum MIUR



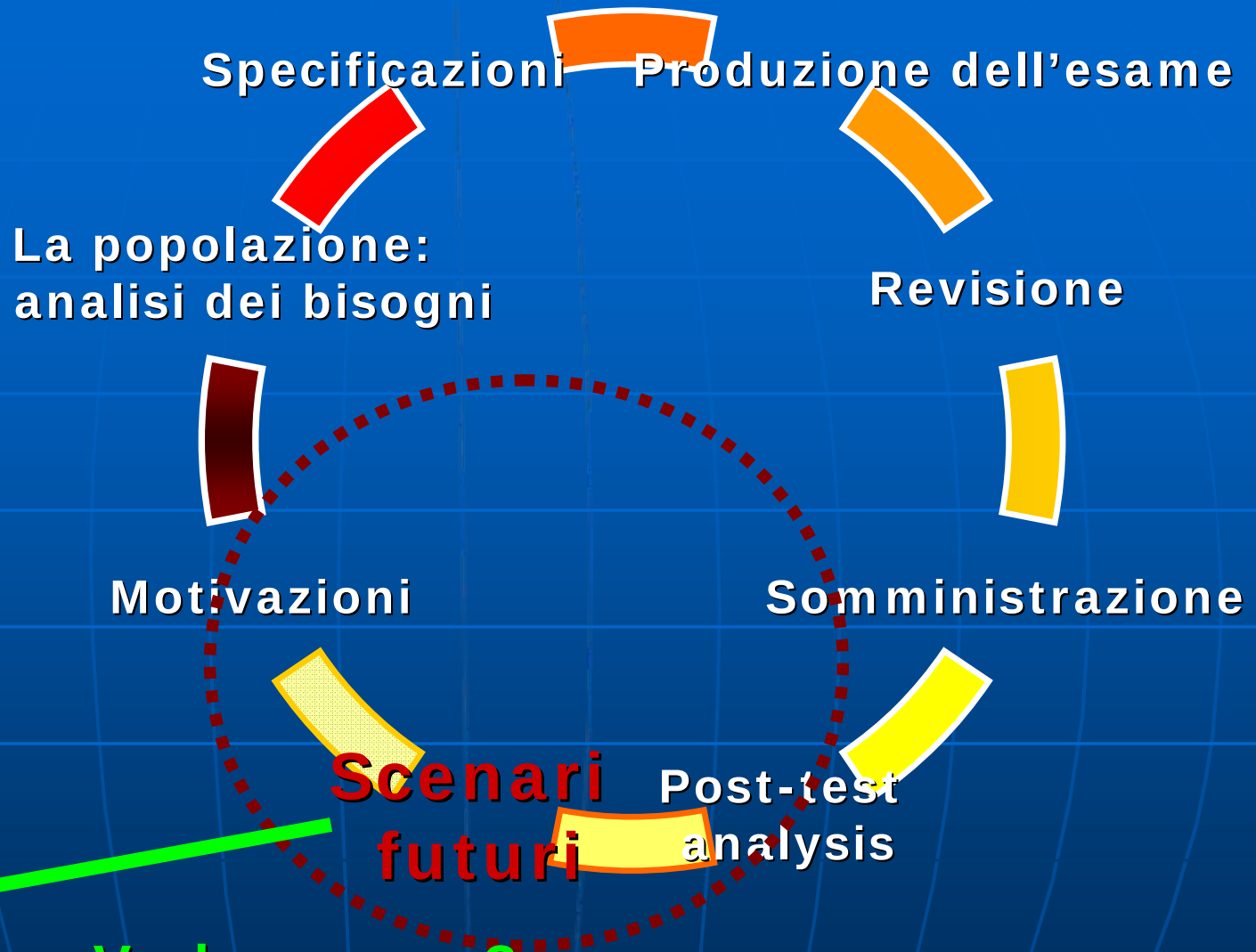
1. Commissioning (reclutamento item writers)
2. Elaborazione prove







- Sessioni d'esame
- Validazione del percorso



- Revisione Vademecum?
- Accordo di Integrazione?
- Nuove tappe ATS?
- KoS test?

Dove lavoreremo soprattutto ...



Obiettivi della formazione

1. Chiarimenti in merito alla normativa
2. Linee guida procedurali per la revisione del test e per la validazione del percorso
3. Tavola rotonda sui possibili scenari futuri
4. **Trovare un provvisorio (delicato, forzato) compromesso fra l'anomalia delle doppie Specificazioni**
5. **Allenarci nella costruzione del test**

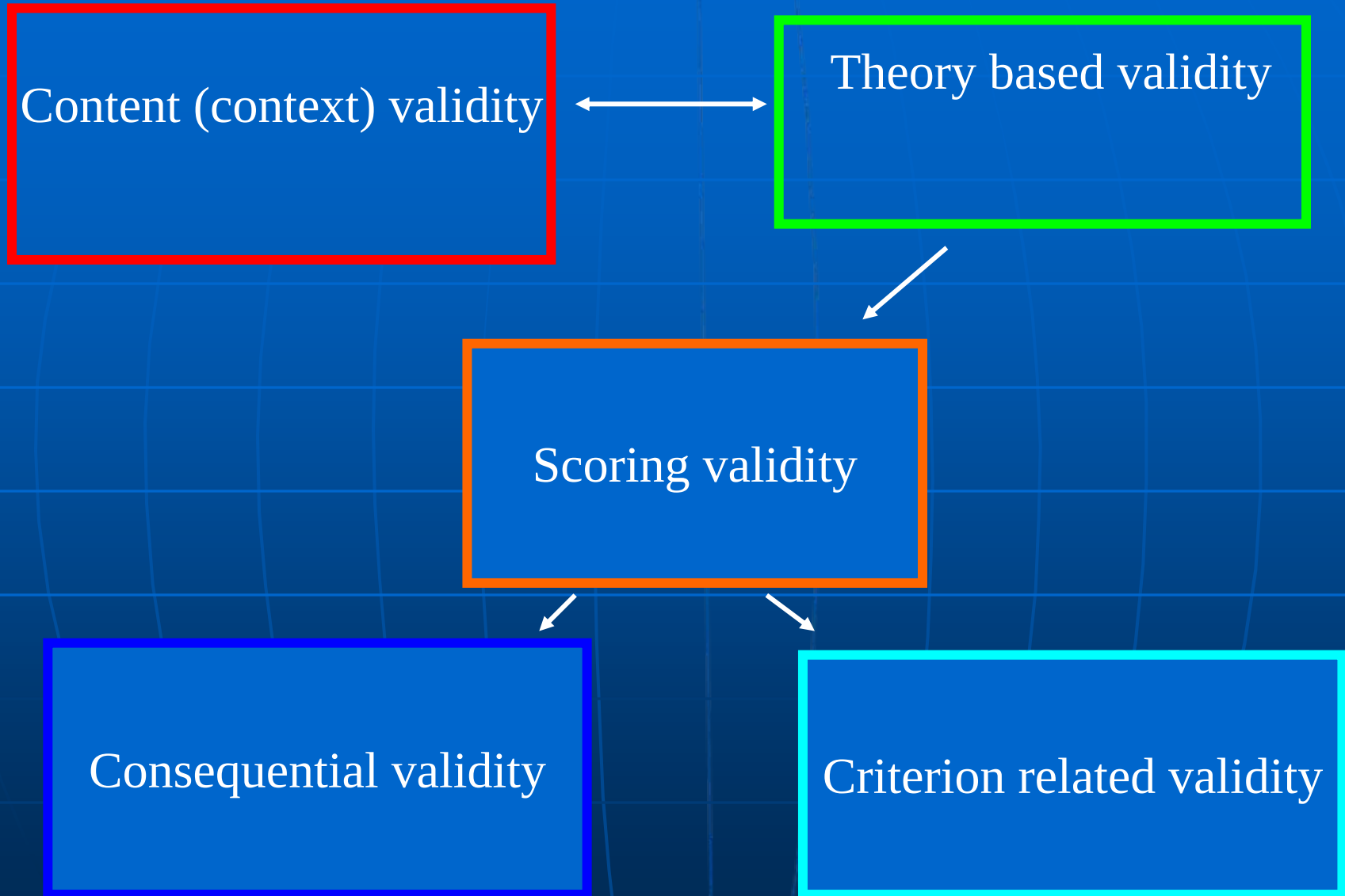
Agenda dei lavori

- Concetti chiave alla base della costruzione di un test di competenza
- Aspetti procedurali e linee guida
- Livello A2 e Sillabo ATS
- Analisi del Vademecum: possibili soluzioni
- Impianto consigliato del test
- Elaborazione delle prove:
 - Lettura
 - Scrittura
 - Ascolto
 - Orale
- Quali scenari?

Tre concetti fondamentali

- **Validità**
- **Affidabilità**
- **Praticabilità**

I 5 volti della validità nell'analisi a posteriori



Content (context) validity

- I contenuti ed i contesti sono vicini all'agire quotidiano dei candidati?
- C'è legame fra la lingua dell'esame e la lingua della vita? C'è la supposta corrispondenza fra *language test performance and language use*?
- È stata condotta in maniera adeguata l'analisi dei bisogni dell'esaminando?

Theory based validity

- Le scelte compiute in sede di elaborazione delle prove sono giustificabili sul piano degli approcci metodologici assunti (costrutto cognitivo)?
- La prova va ad indagare ciò che ci era prefissati di voler indagare? Sono state eliminate le informazioni parassite?

(Hughes)

Scoring validity

- Gli item funzionano? Discriminano?
Sono troppo facili? Sono troppo difficili? Sono ambigui? Sono utili?

Criterion related validity

- Rispetto ad altre sessioni del passato, com'è l'esame?
- Rispetto ad altri test A2 previsti dalla normativa vigente e somministrati in altre Province, come si relaziona l'esame?

Consequential validity

Impact

- Conseguenze del test per il candidato, per la classe e per la società

Correttezza ed eticità del test

(Hessik)

Washback

- Conseguenze del test per l'insegnamento
- Relazione fra formazione e valutazione

Language testing in ottica interdisciplinare

(CoP EALTA)

Cfr. codici ILTA, ALTE, EALTA

Potenziali criticità

- Conflitto fra orientamenti politici ed utilizzo distorto dei certificati linguistici



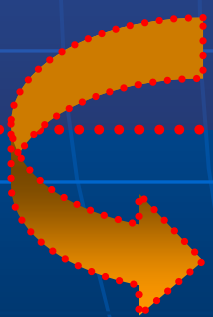
rischio

The diagram consists of a blue rectangular box with a black border containing the word 'rischio'. A large, hollow, downward-pointing arrow is positioned below the box, pointing towards a large, light-blue oval. Inside the oval, there is text describing a shift in testing focus.

da test di competenza
a test di contenimento / controllo dei flussi

Responsabilità del processo valutativo legato a inclusione sociale

- La dimensione etica del test
- Le ripercussioni politiche dell'esito di un test



Mettere in atto rigorose procedure di **buona prassi**
nella costruzione di certificati linguistici
rimane fondamentale ma NON è più sufficiente

Bisogna riflettere sull'**impatto**
del test a livello sociale

Language testers, non tecnici
(T.McNamara)
The power of test
(E.Shohamy)

Il panorama europeo

- **75 % Paesi CoE**

prevede corsi e test
per “migration
purposes”



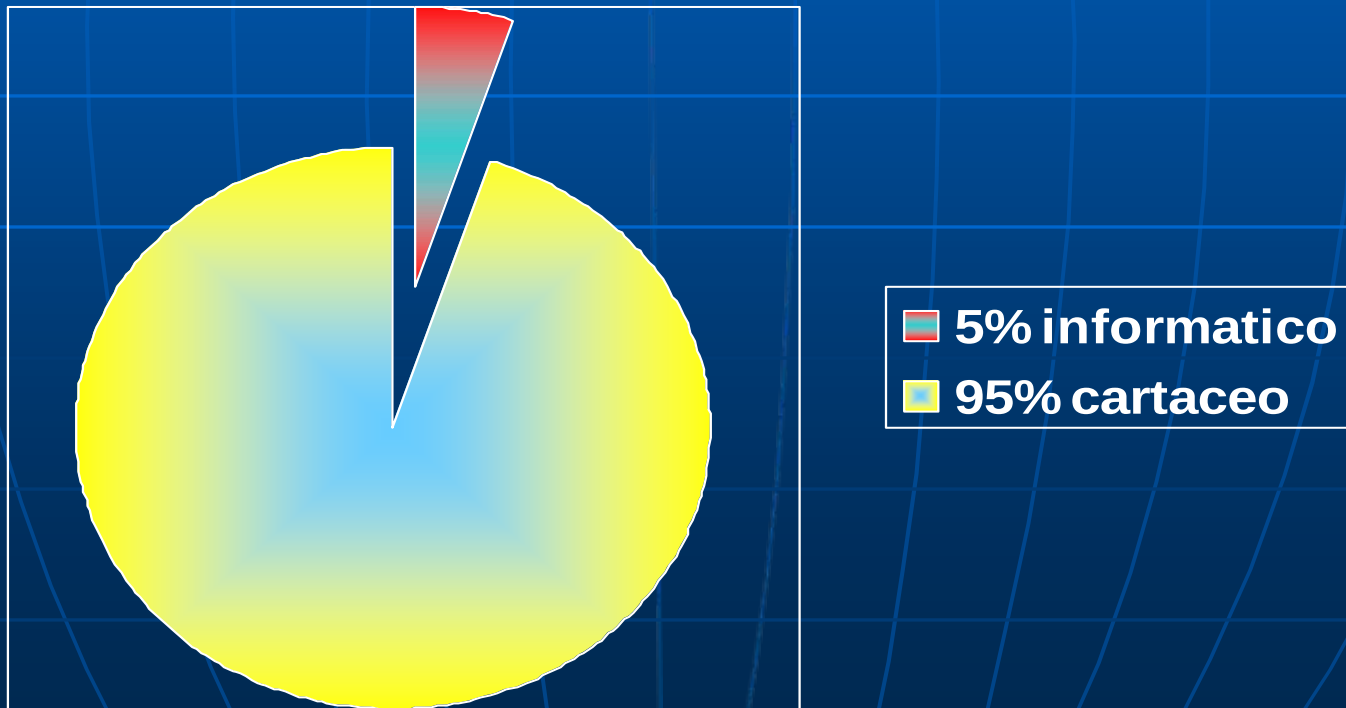
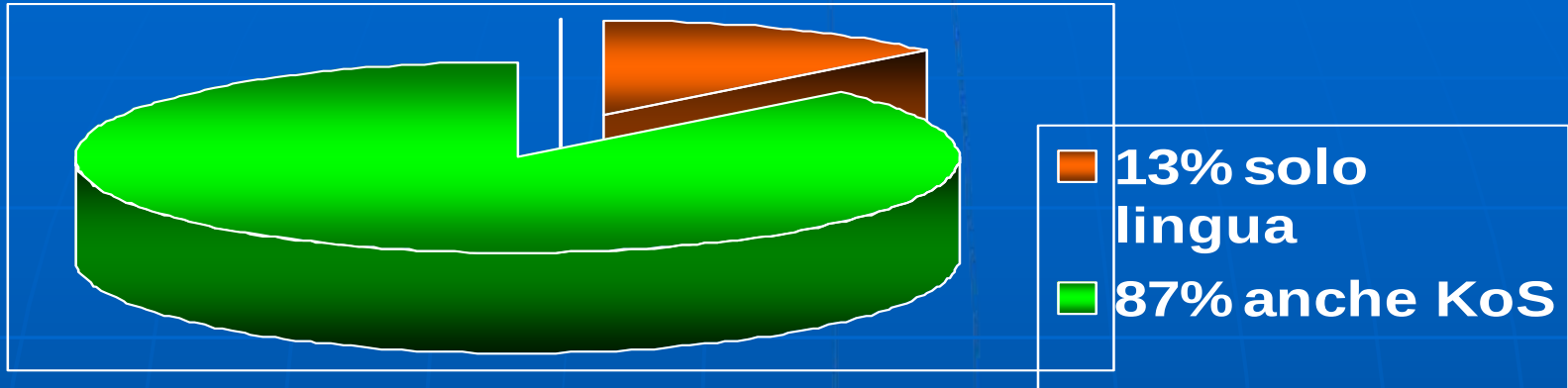
- Ma solo nel 38% con
specifiche linee guida
e percorsi dedicati

**L'A2 è il
livello
mediamente
richiesto**

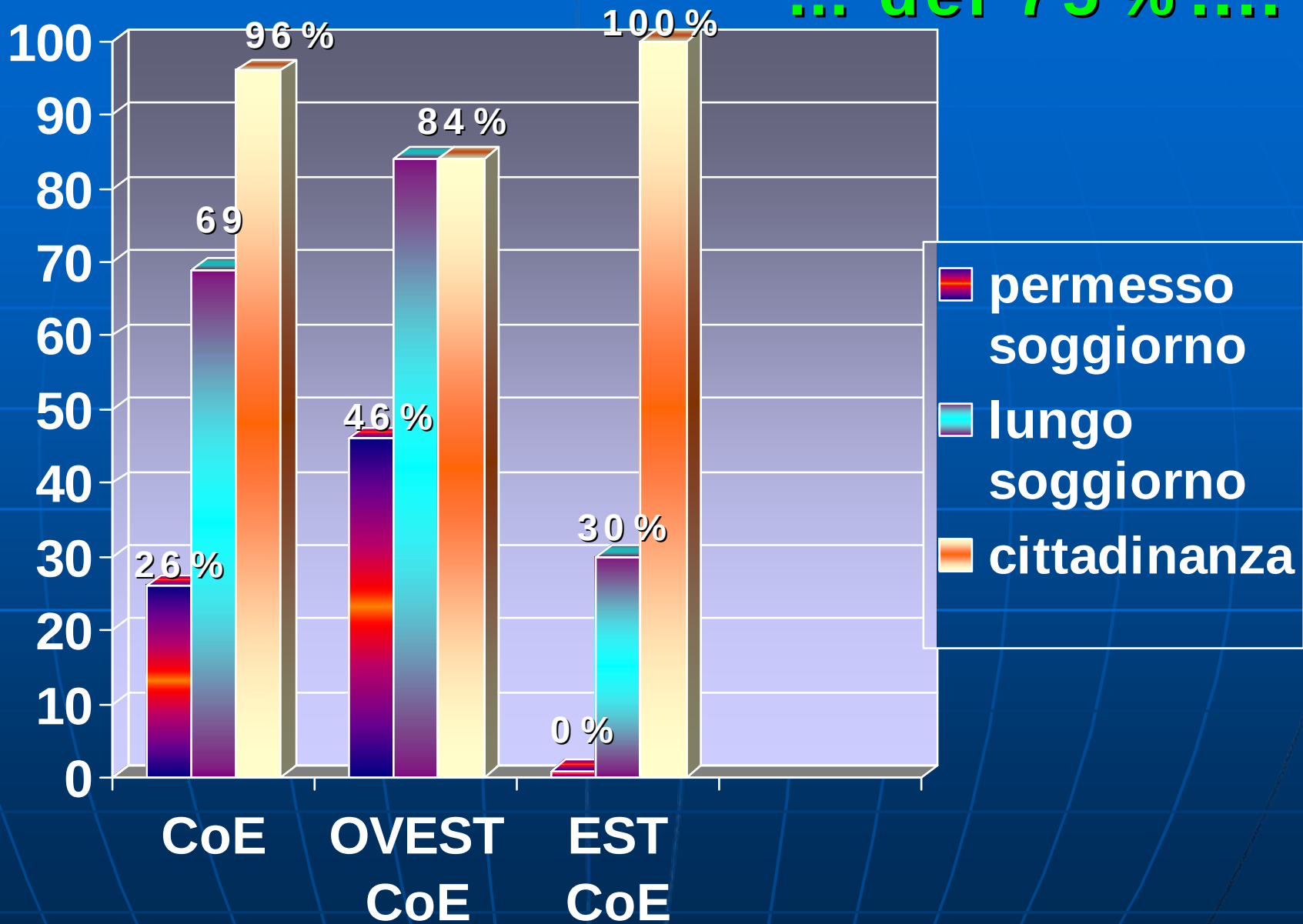


(Van Avermaet: 2010)

... del 75 %



... del 75 %



Agenda CoE: 8 punti chiave

1. Analisi dei bisogni
2. Combattere drop-out
3. Lavorare sulla motivazione
4. Creare percorsi ad hoc
5. Formazione specifica insegnanti
6. Costruire test ad hoc
7. Assicurare adeguate procedure di controllo della qualità
8. Studiare l'impatto



Fairness nel language testing

...per non discriminare...

- **Standard – based approach** (USA 1999), al primo posto:

Garantire equità nell'opportunità di apprendimento

(legame fra insegnamento e valutazione)

- **Test fairness framework** (Kunnan: 2004), poggia su 2 principi:
 1. **Principio di giustizia**: stessi risultati per persone con stesse abilità
 2. **Principio di beneficio**: il test dovrebbe portare beneficio all'intera società

Contenuti e codici etici

- Per un test “fair”:

NO

Guerra, morte, gravi lesioni, abusi, violenza, malattia, politica, religione, sesso, visioni sessiste o classiste, divise ed eserciti, peccati capitali, disastri naturali, imperialismo e colonialismo, stereotipi, visioni etnocentriche, posizioni connotate sul piano dialettico, approcci sarcastici ecc.

Affidabilità

- L'affidabilità è un concetto tecnico relativo alla stabilità dei risultati sia all'interno della medesima prova sia attraverso somministrazioni successive
(Hughes, Bachman e Palmer)
- Una prova quindi è tanto più affidabile quanto più garantisce risultati oggettivi, ripetibili e quindi attendibili

affidabilità → **oggettività** → **equità**

Sdoppiarsi del concetto di affidabilità

interna

esterna

- Afferente all'analisi statistico-quantitativa, al pre- testing (psicometria)
 - Obiettivo nel **breve termine**: assicurare il corretto funzionamento item: strettamente collegata alla scoring validity (calcolata attraverso vari indici: difficoltà, discriminazione ecc.)
 - 2 obiettivi nel **medio periodo**:
 1. creazione banca item
 2. Inserimento nei test di anchor item
- 
- Relativa alla **formazione di TUTTI** i soggetti coinvolti:
 - Istituzioni e dirigenti
 - Personale amministrativo
 - Item writers
 - Insegnanti
 - Somministratori della prova
 - Esaminatori
 - Valutatori della prova orale
 - Correttori delle prove scritte

Validità ed Affidabilità

- Giusto “mix” tra prove oggettive e soggettive, metodi diretti ed indiretti.
- Davies parla di “tensione fra V e A che deve sfociare in equilibrato bilanciamento”, in un accettabile grado di *entrambe* nella VERIFICA



Praticabilità

E' relativa ad aspetti pratici ed organizzativi (fattibilità logistica):

- sia in fase di progettazione
- sia in fase di somministrazione

Alcuni esempi:

- l'acustica per la prova di ascolto
- la sistemazione spaziale dei candidati

Bibliografia ragionata