
Demystifying item writing: The need for a theoretical framework

Xuan Minh Ngo


VNU University of Languages and International
Studies

ngoxuanminhulisvnu@gmail.com

ngoxuanminh@vnu.edu.vn



Presentation Highlights

1. Why item writing?
 2. CHAT in brief
 3. An illustrative case study
 4. The way forward?
- 

1. Why item writing?

- “...one of the critical phases of test development; however, literature on item writing has been sparse” (Kim et al., 2010, p.160) (also Green & Hawkey, 2011; Shin, 2012)
- Conflicting views of item writing
 - Creative arts
 - Realisation of guidelines

1.2. Related studies

- Peirce (Norton) (1992)
 - TOEFL reading
 - Author = test developer
 - “ETS model” (also ETS, n.d.)

**Role of
writers vs.
developer**

External
writers

ETS content
review (test
developers)

ETS stylistic
review

ETS Fairness
Review

1.2. Related studies

- Salisbury (2005), Green & Hawkey (2011)
 - Cambridge Listening + IELTS Reading (AC)
 - Item writing process: 3 phases
 - Collective process
 - Strategies
- ‘non-formalized specifications’”



Figure 1. The Question Paper Production Process for IELTS

1.2. Related studies

- Ingham (2008)

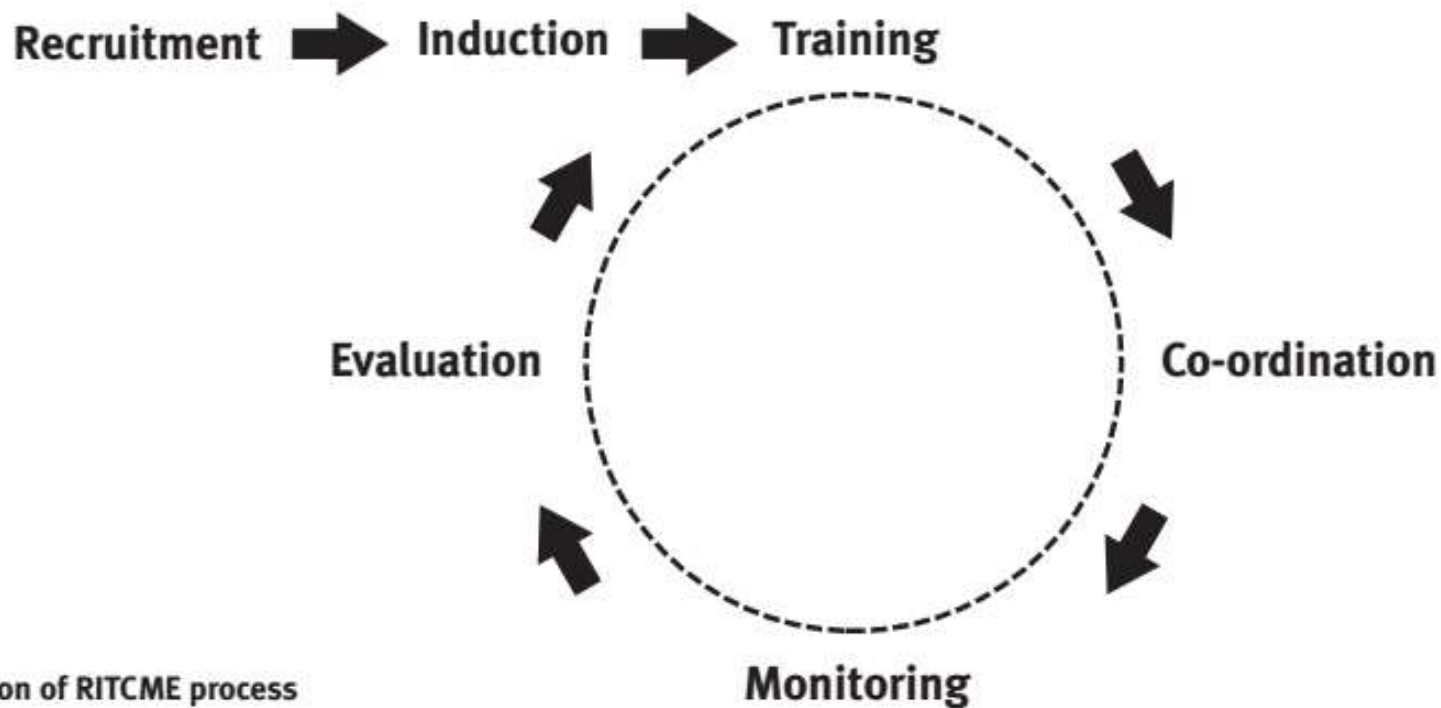


Figure 1: Illustration of RITCME process

1.2. Related studies

- Kim et al. (2010)
 - Practical (experience/ lesson sharing)
 - + Views & use of test specs
 - Involve item writers (IW) + organic guidelines
 - + Group dynamics: personal & collective
 - + Factors: qualifications, experience, personality, background (L/C), preferences

1.3. Gaps

- Remarkable contribution but:
 - Mainly experience sharing (Kim et al. 2010, Peirce, 1992)
 - Lack of coherent analytical framework (except Salisbury (2005))
- Cultural Historical Activity Theory (CHAT):
Activity = “a coherent, stable, relatively long-term endeavor directed to a definite goal object” (Rochelle, 1998)

1.3. Gaps

- Remarkable contribution but:
 - Still highly limited in quantity
 - Different foci: formation of expertise (Salisbury, 2005), text adaptation & authenticity (Green & Hawkey, 201), training (Ingham, 2008)
 - Established, international tests (IELTS, TOEFL) or ESL (Kim et al. 2010)
 - > Homegrown + EFL context

2. CHAT in brief

- What?
- L.S. Vygotsky → Leont'ev & Luria (1920s-30s)
- “mediational roles of tools and artefacts within a cultural-historical context” (Barab, Evans and Baek, 2004, p.204)
- Unit of analysis = a complex human activity

2.1.CHAT 1.0

- overcame Cartesian dualism + “crisis” in psychology
- But
Individual vs. Group?

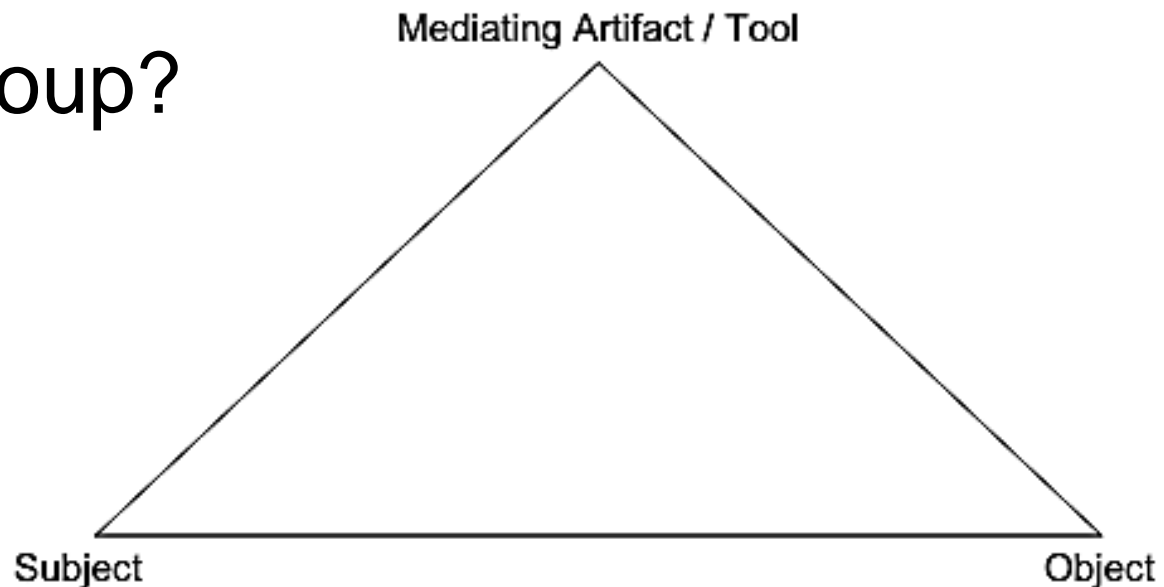
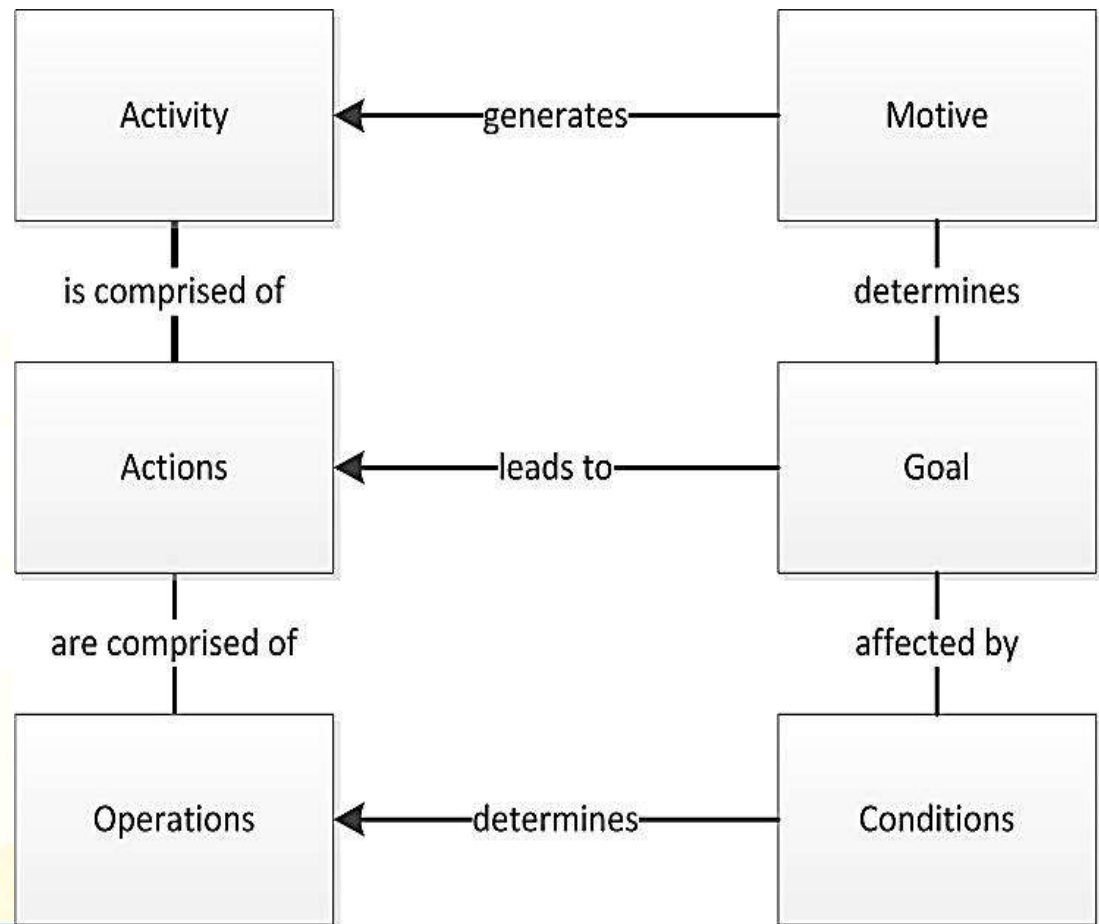


Fig. 1. Vygotsky's basic mediation triangle adapted from Cole (1996).

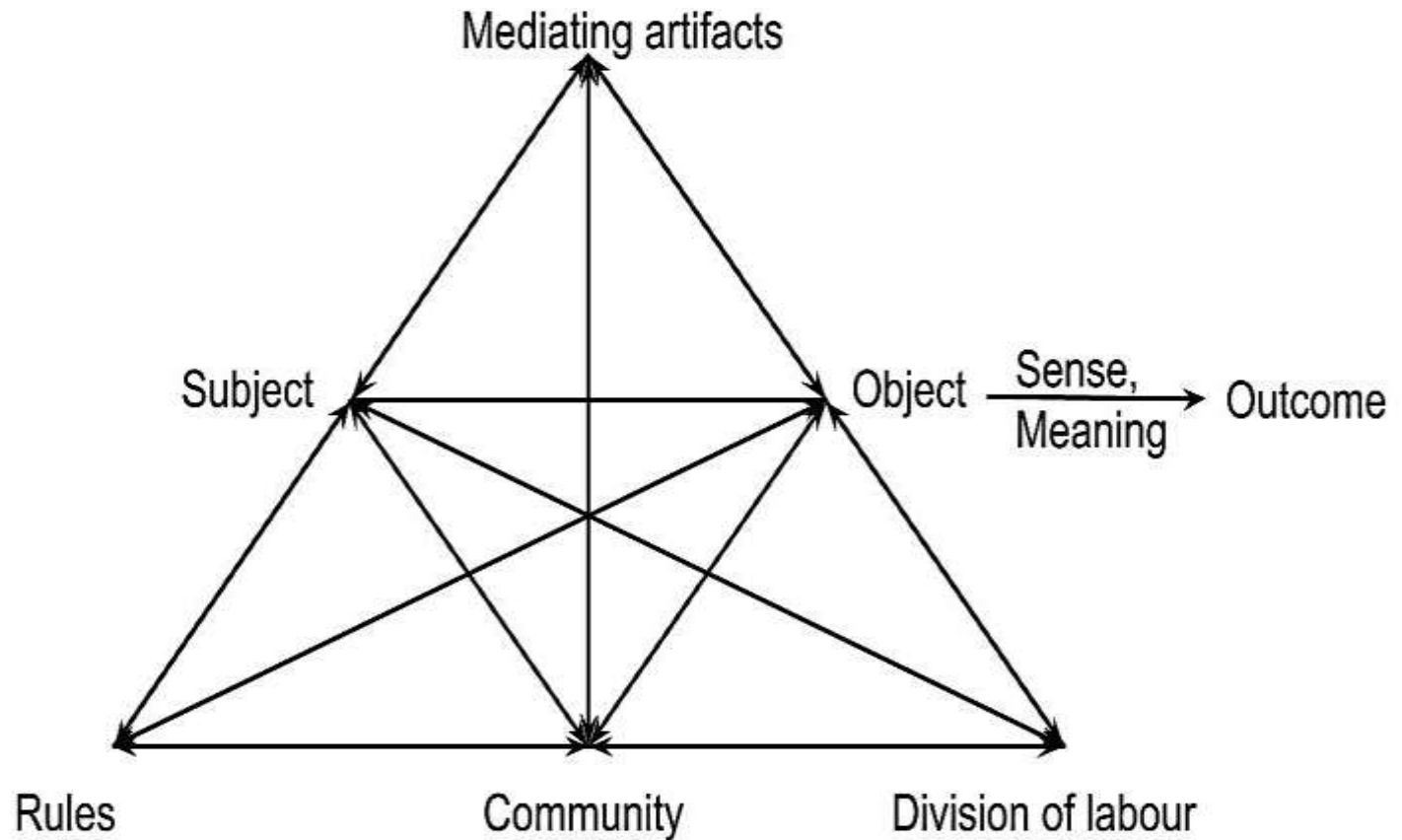
2.2. CHAT 2.0

- Individual action vs. Group activity (Leont'ev, 1981)



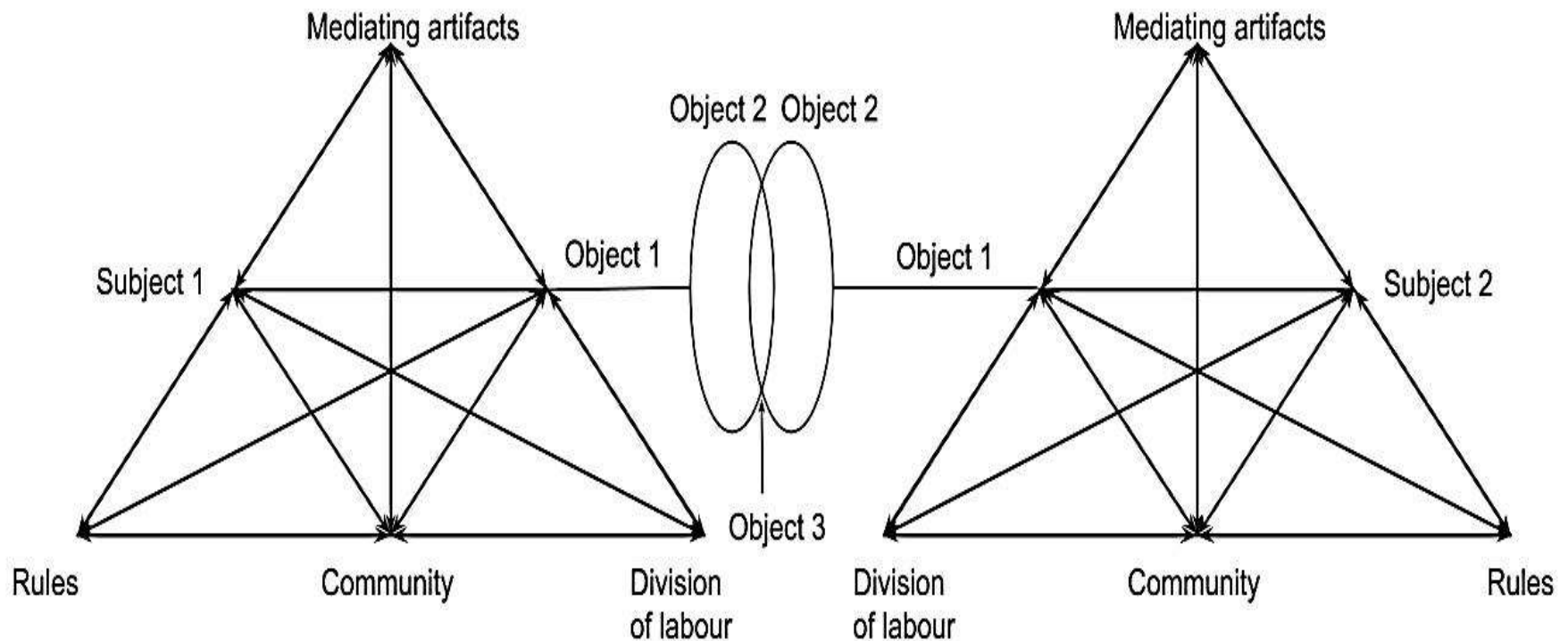
2.2. CHAT 2.0

- Engestrom (1987)



2.3.CHAT 3.0

- Two interacting activity systems (Engestrom, 1999, 2001)




2.4. CHAT & Item writing

- Why?
 - Item writing = complex human activity
 - CHAT
 - + Analytical power (esp. historical development & contradictions)
 - + Design change
 - + Ease of communication (graphic)
- See: Engestrom & Miettinen (1999), Martin & Peim (2009), Roth & Lee (2007), Yamagata-Lynch (2010)

3. An illustrative case study

Part of an ongoing project (Ngo, in preparation)

- Settings
 - Participant
 - Research Questions
 - Methods
 - Preliminary findings & Discussion
- 

3.1. Settings

- Vietnam: EFL / National Foreign Languages Project (2020)
- The tests
 - A suite of homegrown English tests
 - Public university (nationally recognised)
 - CEFR aligned (4 skills, multiple levels)

3.2. Participant

- A “successful” listening item writer (?)
- C2 + MA in Applied Linguistics (Australia) (a course on language testing)
- Experience:
 - + Teaching: 7 years (English majors) + IELTS preparation classes
 - + Test development: school (regular) + university (1 project)
 - + Item writing: 15 tests (higher level) + 8 (current test)

3.3. Research Questions

- What are the factors that mediate the item writing activity?

Subject, tools & signs, object, outcome, rules, communities and division of labour

- What are the major contradictions in this activity system?

Contradictions = driving force of change & development (Engestrom, 2001)

3.3. Methods

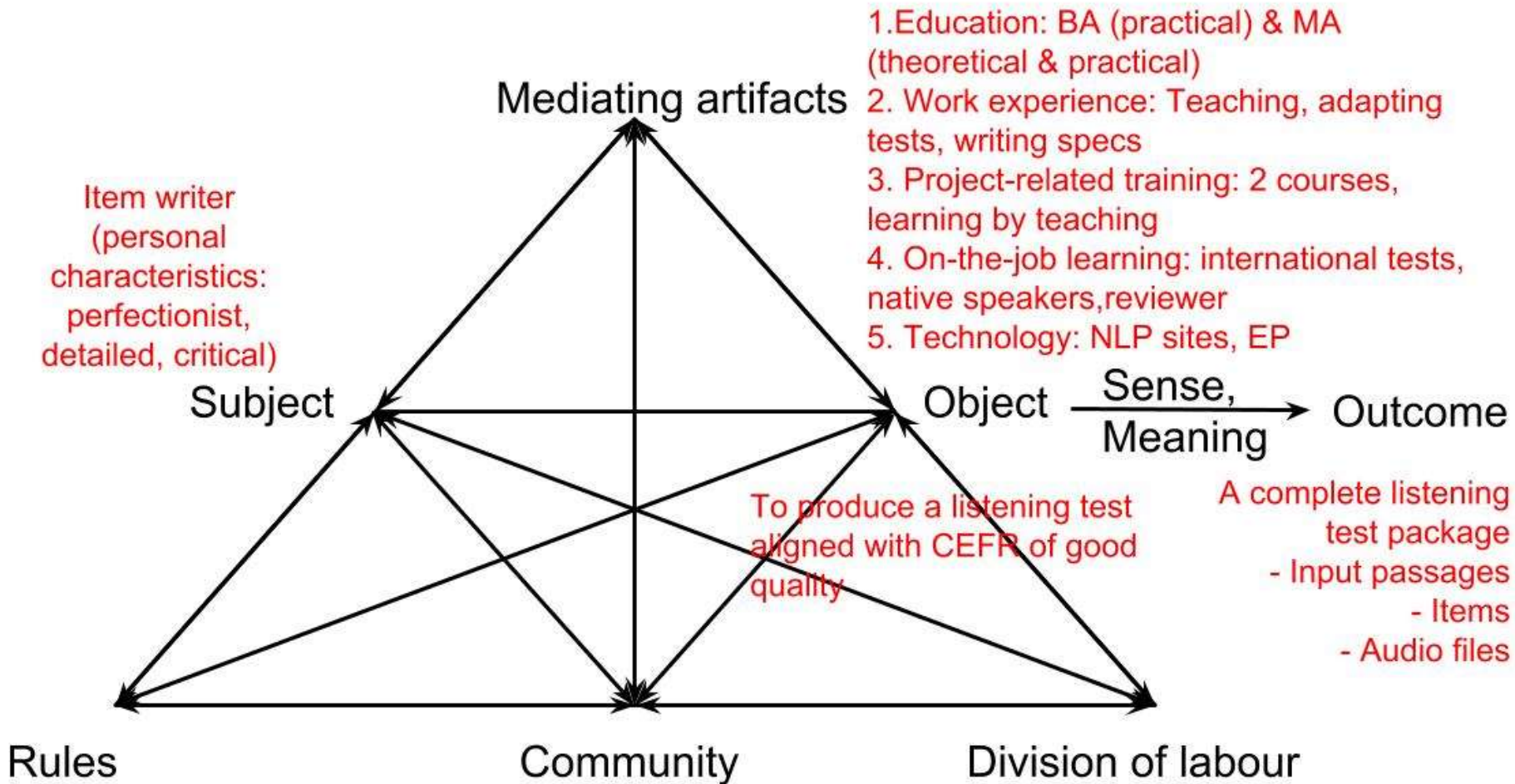
Collection

- Narrative frames (Barkhuizen & Wette, 2008)
- Verbal protocol (McKay, 2009)
- Reflective journals (Nunan, 1992)

Analysis

- Constant comparative method (Glaser & Strauss, 1967; Strauss & Corbin, 1998; Corbin & Strauss, 2008)
 - Open coding
 - Axial coding
 - Selective coding

3.4. Preliminary findings & Discussion



3.4. Preliminary findings & Discussion

- Mediating artifacts - Education
- “I took my Master’s course at the University of X. It was during the Language Testing and Assessment course taught by Dr. Y. One of our major assessment tasks was to design an achievement test based on a textbook unit of our choice, starting from test specifications, then test items, guidelines for stakeholders and a critique of our own test. **It was the very first time that I heard the term “item writing”.**

3.4. Preliminary findings & Discussion

- Mediating artifacts - Education
- Actually, I crafted listening items before during my BA course. The first time was in the second year, semester 1. I paired up with a classmate and we designed a listening mini-test based on a YouTube video consisting of gap filling and short answer questions But then we didn't really use the term "write test items", just "design questions" probably because we were taught by teachers who didn't have a background in language testing. It was useful but somehow I thought it gave me the impression that I could write questions, but only based on a preexisting recording.

3.4. Preliminary findings & Discussion

- Mediating artifacts - On-the-job learning & Technology

“I used some natural language processing websites like **LexTutor and Readability**. Also, the **Vocabulary Profile and Cambridge Dictionaries** are of great help. But I also rely on **Cambridge exam books**. I often do some tests in those to **get a sense of what it means to be at B1, B2 or C1** as well as to get ideas about scenarios for the tests. By the way, I did refer to the CEFR, but **after some time I stopped to**. Now I just internalize some key words like for B1 it should be familiar, concrete and specific? For B2, it's a mix of concrete and abstract. For C1, definitely it must be complex, abstract, unfamiliar. So yes, the CEFR does have a lot of bearing on the way I select topics. But I must say Cambridge books have a great role to play because they realise what the CEFR implies. **And when I write items, I prefer something clear, simple and direct.**”

→ Kim et al. (2010): item writer's preference for clear, straightforward samples

3.4. Preliminary findings & Discussion

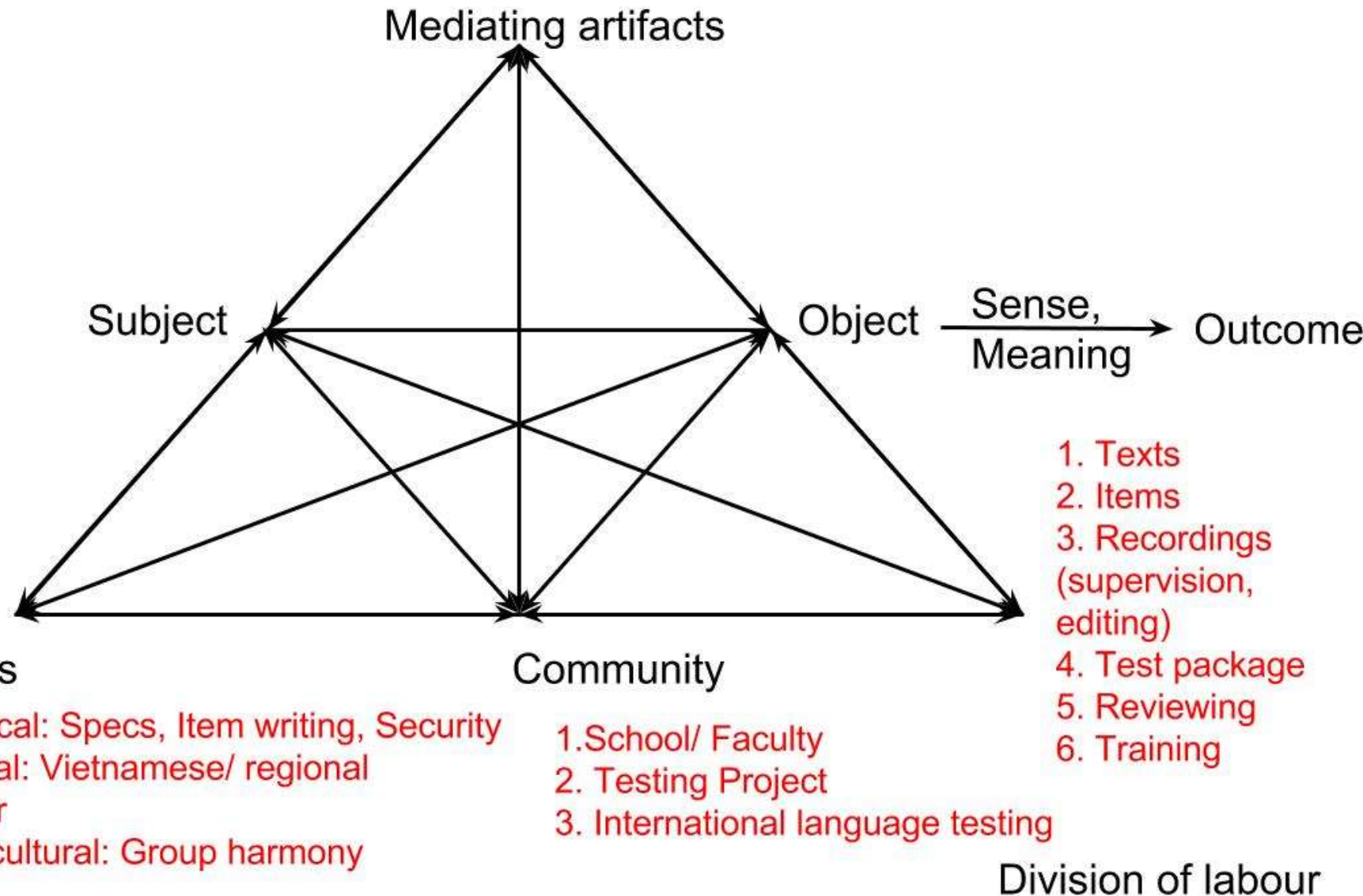
- Mediating artifacts – On-the-job learning

“And we love coming into contact with native speakers. We felt like we learnt so much from them. They help us fix mistakes in our expression.”

→ Native speakers’ role in a test written by non-native speakers?



3.4. Preliminary findings & Discussion



4. The way forward

- Explore the systemic contradictions in this activity system

E.g. Subject >< Object, Subject >< Rules

- Explore the joint activity system

E.g. Item writing as a collective process

→ A group of item writers

→ Item writing interacting with other activities (e.g. Item reviewing/ editing)

4. The way forward

- Interventionist studies
 - Involve different stakeholders: item writers, administrators, reviewers, etc.
 - Group discussion based on the activity system analysis to resolve contradictions
- Change laboratory (Engestrom et al., 1996)

A final word!



Item writing: crucial but under-research

CHAT: a analytical framework

CHAT: an interventionist tool

THANK YOU FOR YOUR ATTENTION

A decorative background featuring a grid of interlocking puzzle pieces in shades of red, orange, and yellow, set against a light, hazy background. The puzzle pieces are arranged in a pattern that suggests a path or a solution, with some pieces missing or slightly offset. The overall effect is one of complexity and problem-solving.

Q & A

