

‘Process Data and its Uses’

CEMO Lecture Series on *Process Data in Educational Assessment*

Lecture 1, January 9th, 2023.

Dr Bryan Maddox



UiO : **University of Oslo**



**Assessment
MicroAnalytics™**



DIGITAL
EDUCATION
FUTURES
INITIATIVE
CAMBRIDGE



Monday 9th: Process Data and Its Uses

Wednesday 11th : Test Design and Administration + Observation Activity

Thursday 12th : Extending analysis of Performance and Engagement, and Key Reading Discussion (Ercikan, Guo and He)

Friday, 13th : Eye Tracking + Practical activity with TA and RTA. Discussion of Module Assessment.

Monday 16th : Process Data and Validity, and Key Reading Discussion 2 (Goldhammer et al).

Tuesday 17th : Homework Task – Preparation for Data Collection.

Wednesday 18th : Data Collection & Coaching, Preparation of Presentations.

Thursday 19th : Presentation, Formative Assessment Task, Module Evaluation.

Statement 1

‘Occasionally the product of performance is *not* measurable apart from the performance. Playing a musical instrument is an example of such a case.... In other instances, *both* the performance in process and the product of the performance may be measured by judging an individual’s proficiency at a given sort of behaviour... it is highly desirable to measure actual operation, or performance in process in many cases..’

Ryans and Frederiksen, 1951, Cited in Ercikan, Guo and He 2020, p2.

Statement 2

‘Examinees’ final answers to assessment tasks, “responses” or “products” may not provide information about what kind of response processes examinees engaged in to reach their responses, it may not be possible to determine whether examinees engaged with the tasks in the intended ways, and whether different examinees took different solution strategies to produce the same responses’.

(Ercikan, Guo and He, 2020, p1).

Statement 3

‘The increasing use of computer-based testing and learning environments is leading to a significant reform on the traditional form of measurement, with tremendous extra available data collected during the process of learning and assessment (Bennett et al., 2007, 2010). It means that we can learn and describe the respondents’ performances not only by their responses, but also their responding processes, in addition to the response accuracy in the traditional tests (Ercikan and Pellegrino, 2017).’

Jiao, He and Veldkamp, 2021. p1.

Statement 4

‘In the standard assessment paradigm, the application of evidence identification rules means scoring the work product created by the test-taker. However, when “scoring” the work process (i.e., creating a process indicator), evidence identification refers instead to identifying the presence or absence of certain low-level features (i.e., actions or states) or certain patterns of low-level features in the continuous stream of log events elicited during the item completion process. Accordingly, evidence identification in the traditional sense is scoring a time-bounded work product, whereas in continuous assessment, it refers to feature extraction and aggregation based on continuous behavior over time (Behrens & DiCerbo, 2014).’

Goldhammer et al. (2021) ‘From byproduct to design factor: on validating the interpretation of process indicators based on log data’, Large-scale Assess Education (2021) 9:20

Statement 5

‘Informed by Cronbach and Meehl’s description [of test constructs], we consider that the attribute(s) of the test taker about which we make claims in interpreting a test include:

- ‘Test-construct(s),’ reflected in the product-oriented test-taker response data to the items or tasks in the test (that is, their deliberate and conscious responses), and;
- ‘Process-construct(s),’ reflected in the process data recorded by the various sensors and stored in the log files, including keystroke data, eye-tracking, and the associated time stamps, performed by a test-taker as they complete a task.’

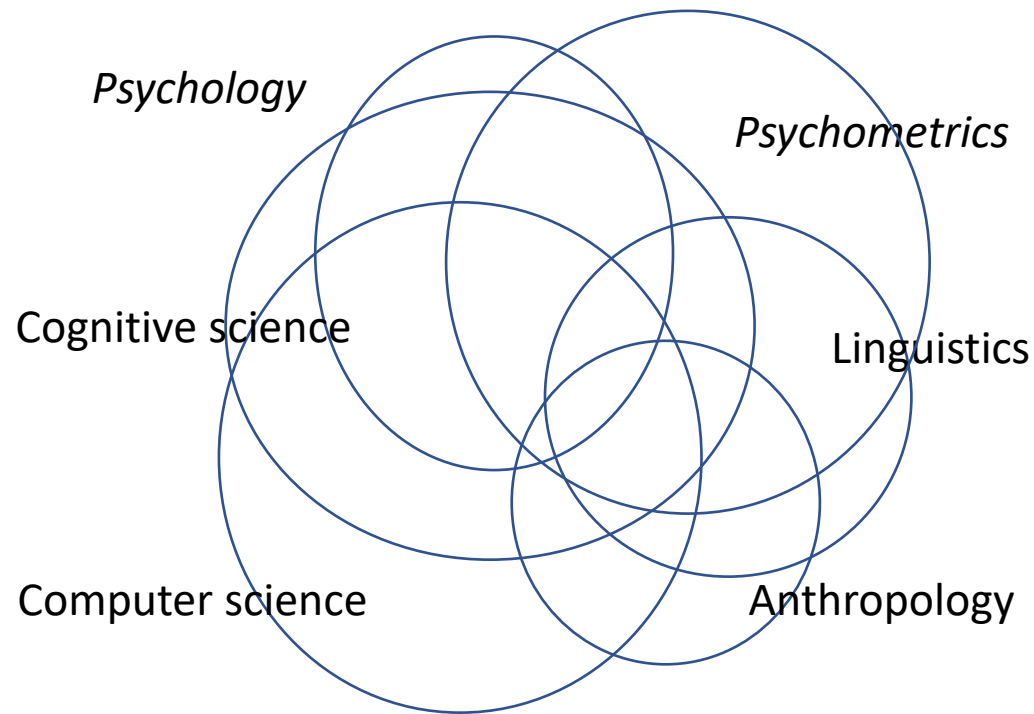
Zumbo, Maddox, and Care (in press, 2023). ‘Process and Product in Computer-Based Assessments: Clearing the Ground for a Holistic Validity Framework. *European Journal of Psychological Assessment*’.

Discussion Question

What do each of these statements tell us about the way that the authors understand assessment response processes and their uses?

Task: Please discuss the statements and take some notes, comments, questions on each one.

Assessment Response Processes: A Field of Enquiry



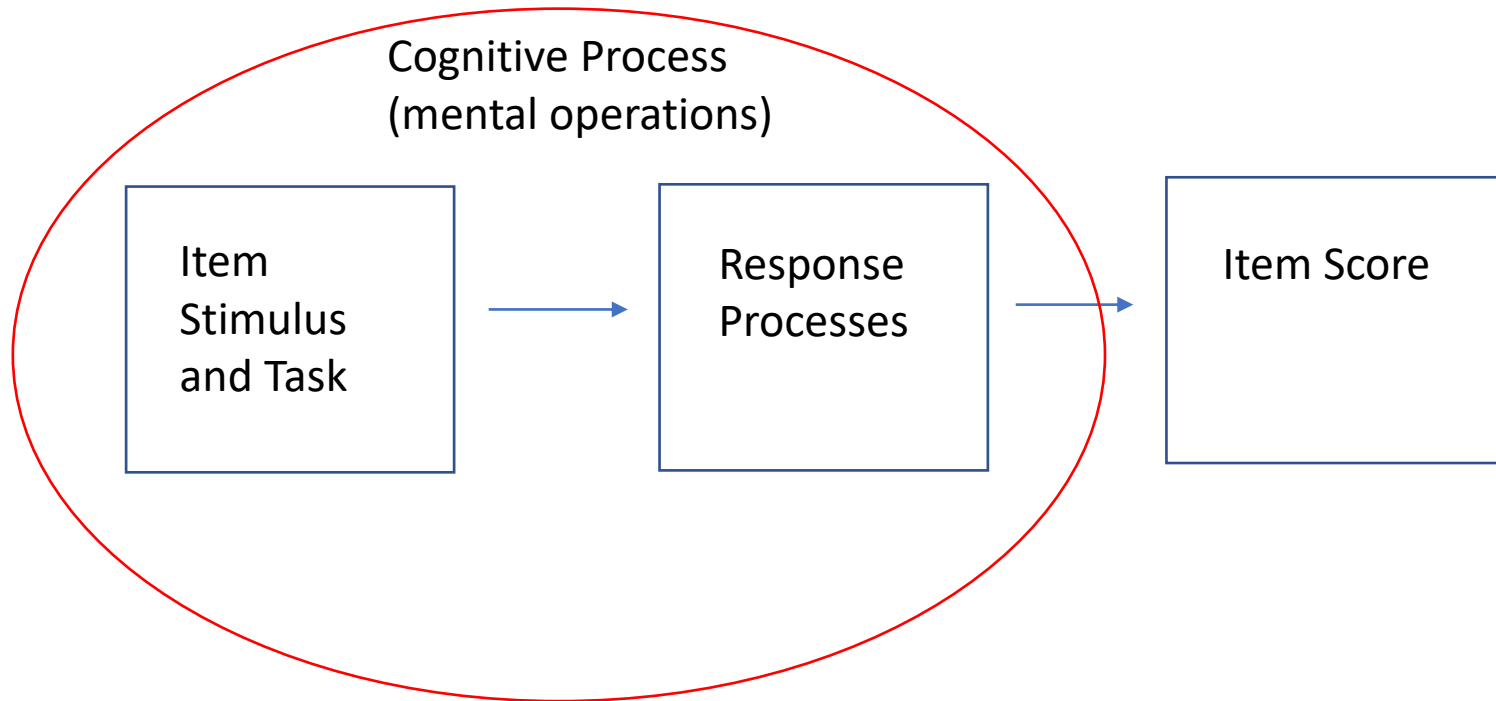
Defining Response Processes (1)

‘Response processes refer to the thought processes, strategies, approaches, and behaviours of examinees when they read, interpret and formulate solutions to assessment tasks.’

(Ercikan and Pellegrino, 2017, p2).

Discussion Questions

1. Which of the dimensions of response processes described by Ercikan and Pellegrino (2017) can be observed or measured?
2. Which of those dimensions might help some aspect of educational assessment?



In 'information processing' model of assessment response processes. e.g. Embretson (2016); Bruckner & Pellegrino (2016); Kane and Mislevy (2017).

Defining Response Processes (2)

'..one may think broadly of response processes as the mechanisms that underlie what people do, think or feel when interacting with, and responding to, the item or task and are responsible for generating observed test score variation.'

(Huble and Zumbo, 2017, p2).

Discussion Question

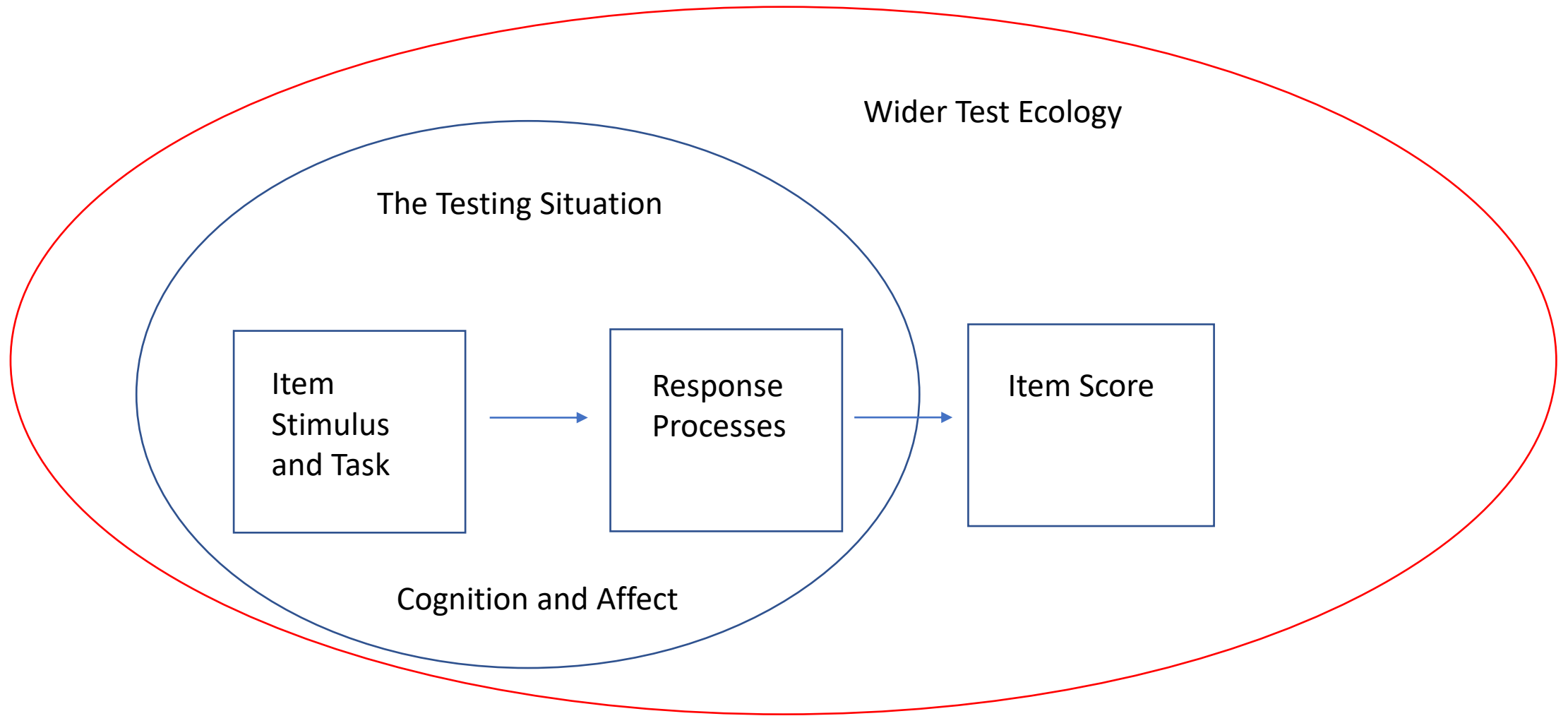
What does the Hubley and Zumbo (2017) add to the definition of assessment response processes proposed by Ercikan and Pellegrino?

For a review and discussion of both books, see Newton, P.E. (2019) 'What is response process validation evidence and how important is it? An essay reviewing Ercikan and Pellegrino (2017) and Zumbo and Hubley (2017)', *Assessment in Education: Principles, Policy & Practice*, 26 (2), 245-253.

Beyond the cognitive realm

‘This definition expands response process beyond the cognitive realm to include emotions, motivations, and behaviors. Inclusion of affect and motives allows us to take into account how these may impact on the different respondents’ interaction with the item(s), test, and testing situation’

(Huble and Zumbo, 2017, p2-3).



An ecological model of response processes. e.g. See Zumbo, B.D. (2017). 'Trending Away From Routine Procedures, Towards an Ecologically Informed 'In Vivo' View of Validation Practices'. *Measurement: Interdisciplinary Research and Perspectives*, 15:3-4, 137-139.

Zumbo, B.D., Liu, Y., Wu, A.D., Shear, B.R., Astivia, O.L.O. & Ark, T.K. (2015). A Methodology for Zumbo's Third Generation DIF Analyses and the Ecology of Item Responding. *Language Assessment Quarterly*, 12, 136-151.

Hidden and Revealed

‘... even in digital assessments, cognitive responses themselves are **not observable**. What is captured in the think aloud protocols as well as in log files need to be considered as “traces of processes” rather than processes themselves.’ (Ercikan, Guo and He, p3, emphasis mine).

‘We argue, however, that one may think broadly of response processes as the mechanisms that ***underlie*** what people do, think, or feel when interacting with, and responding to, the item or task and are responsible for generating observed test score variation.’ (Huble and Zumbo, 2017, p2, emphasis mine).



Observing Response Processes

Respondent: *These graphics are confusing.* (Some visitors are arriving to look at horses that the test taker intended to sell)

Bystander: It's getting dark.

Tester: *Number the answers to make them clear.* (The respondent smiles and concentrates on the task. There is movement of horses outside. A visitor comes in, pours himself some tea and lights a pipe)

Respondent: *I'm just looking at this* (indicating to the visitor the test booklet as the focus of activity)

Tester: *If you think it is difficult you can skip the question.* (The respondent is really concentrating on the task. Three herders are now sitting watching the assessment event. A fourth herder enters)

Bystander (fourth herder): *The horses are coming in.*

Test taker: *Okay. You go for the cows.*

Extract from Maddox, 2015, p437.

Ethnographic Observation

R: Come here and help me! [to **his daughter**]

I: **You can't help him** [to the respondent's daughter – who now stands beside her father] [...continues in same style.....]

R: ten percent is 800, so it's 7,200.

R's Daughter: Correct.

R's Daughter: Dad – the first question is about the kind of stuff that she ate..

R's Daughter: Can he write it down here? [she asks the interviewer]

I: **Sister – you can't help him! Why not go and watch TV.**

R: Daughter! Come here.

R. What is this..? [..The assessment continues in the same style..].

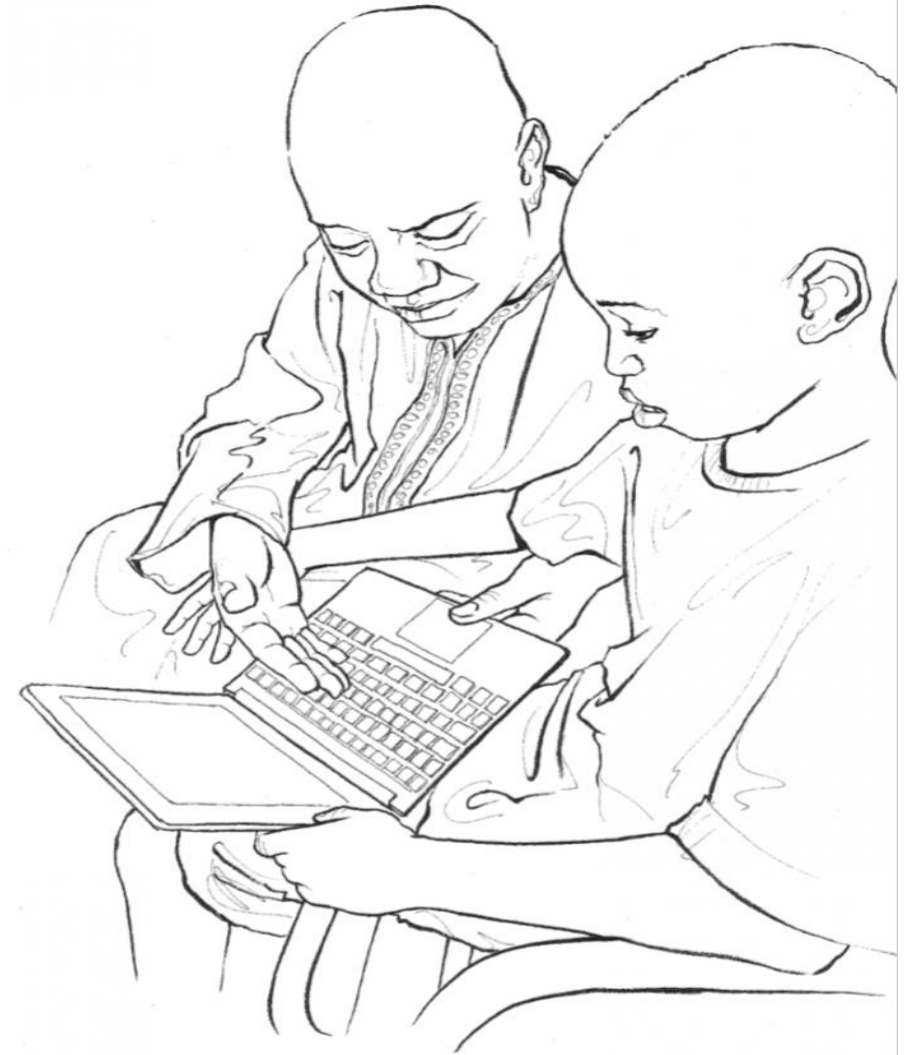
R's Daughter: You should calculate ten percent, and half of that, then add to get fifteen percent. You should get ten percent, and then five percent, and then deduct it from the price.

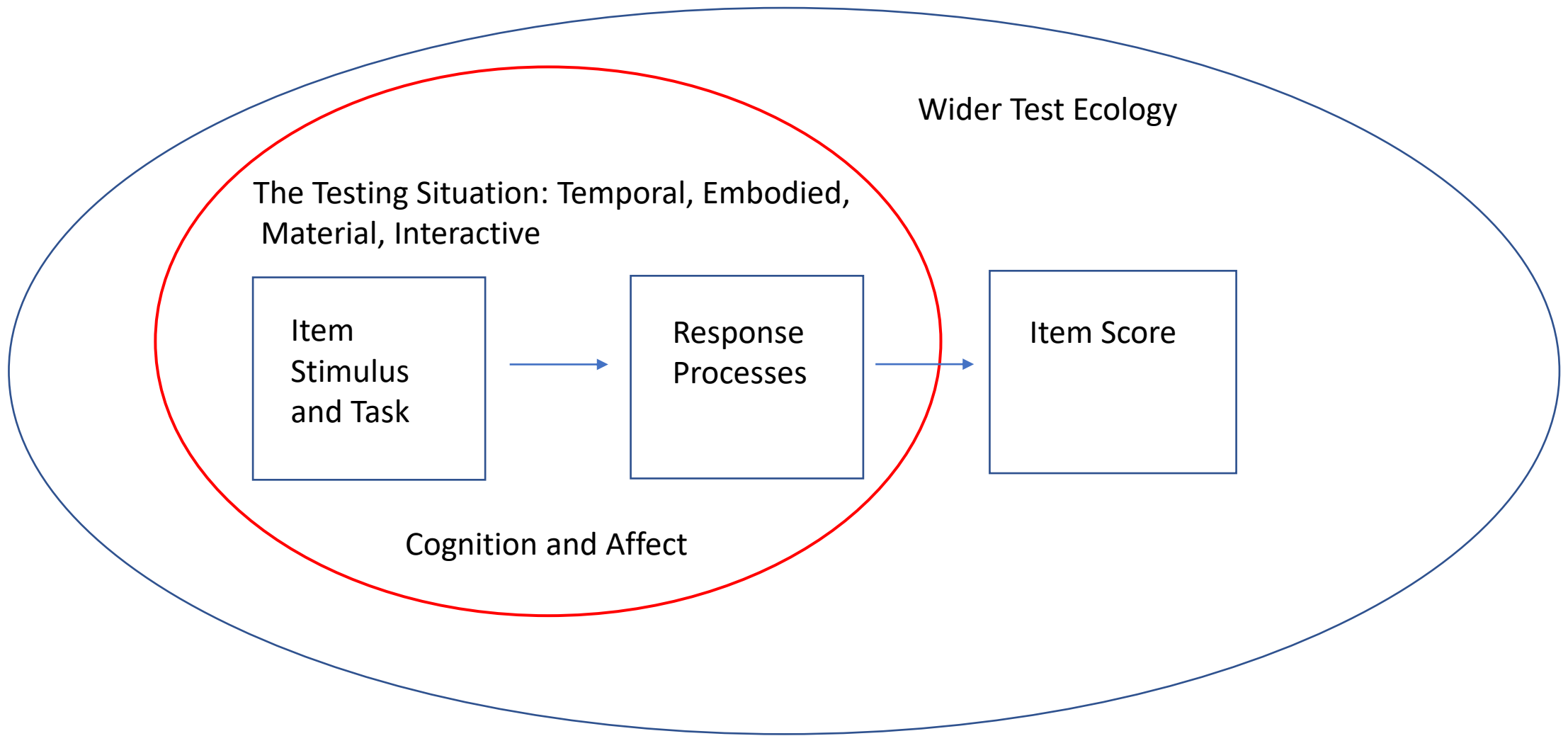
I: **Hey! You should not help him.**

Defining Response Processes (3)

‘Assessment response processes concern observed evidence of the cognitive and affective dimensions of test taking in the testing situation, taking into account their temporal, spatial, embodied, material, interactive and socio-cultural characteristics’.

(Maddox, 2023 – working definition)





Extended, Embodied, Action-Oriented and Environmentally Coupled Cognition
(see Clarke 2011, and Goodwin 2007).

Measurement Opportunities: Sources of Process Data

- Log Data on Response Times, Key Strokes (clickstream).
- Eye Tracking (Screen mounted and Wearable)
- Observation, Video and Audio Recording (ad hoc or systematic, with or without automated transcription and NLP analysis).
- Wearable Sensors (movement, and physiology – heartbeat, GSR, fNIRS)
- Retrospective Think Aloud and Cognitive Interviews (explanations, subjective responses and explanations)

See Oranje et al. 2017; Zumbo and Hubley (Eds) 2017.

U304-ContactEmployer;C304B711;StimulusAndQuestionLoaded;10:50:41

U304-ContactEmployer;C304B711;click;10:50:41

U304-ContactEmployer;C304B711;highlightEvent;10:51:42

U304-ContactEmployer;C304B711;highlightEvent;10:52:21

U304-ContactEmployer;C304B711;highlightEvent;10:52:25

U304-ContactEmployer;C304B711;highlightEvent;10:52:26

U304-ContactEmployer;C304B711;highlightEvent;10:52:28

U304-ContactEmployer;C304B711;highlightEvent;10:52:31

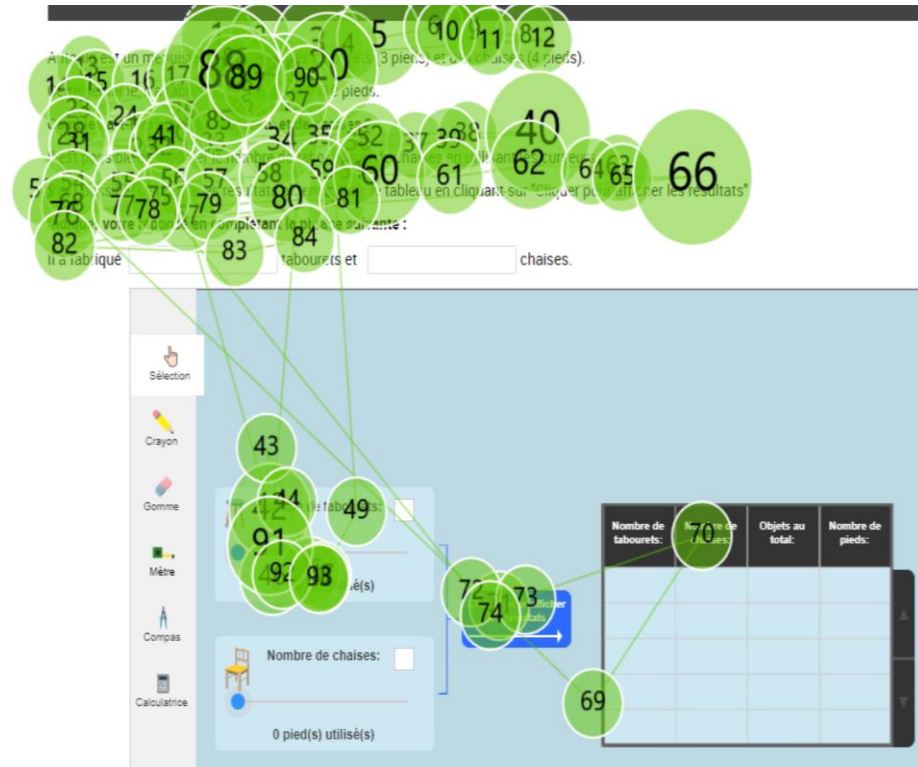
U304-ContactEmployer;C304B711;highlightEvent;10:52:35

U304-ContactEmployer;C304B711;click;10:52:37

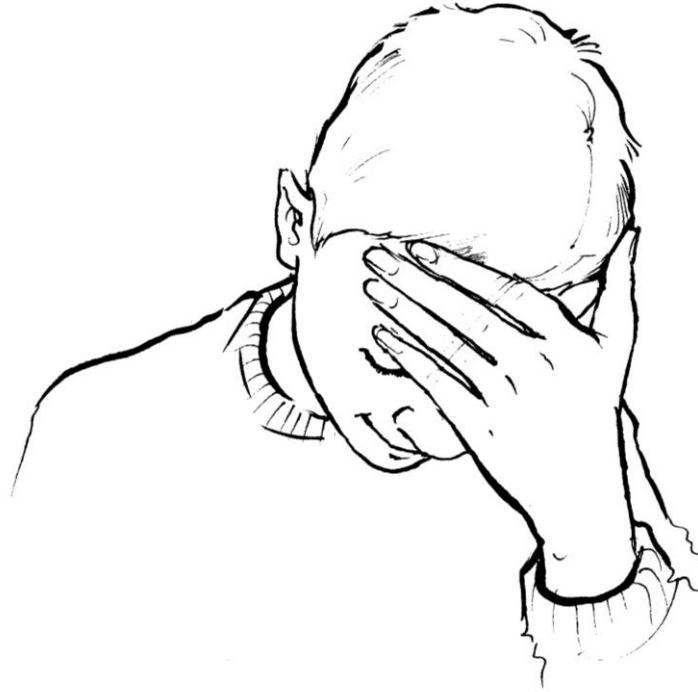
U304-ContactEmployer;C304B711;onItemEnd;10:52:39



Eye Tracking



Video Data on Gesture & Facial Expression



On the use of Gesture and Facial Expression in assessment, see Maddox and Zumbo, 2017.

Think Aloud & Retrospective Think Aloud (RTA)

“..and this one is like 3, so I thought it was that one”.

France Mathematics Tests

B: .. when there was too much text, did you ever give up and skip? Just pressed advance?

R: Um, I think here, in this one ...

B: In that one, yeah.

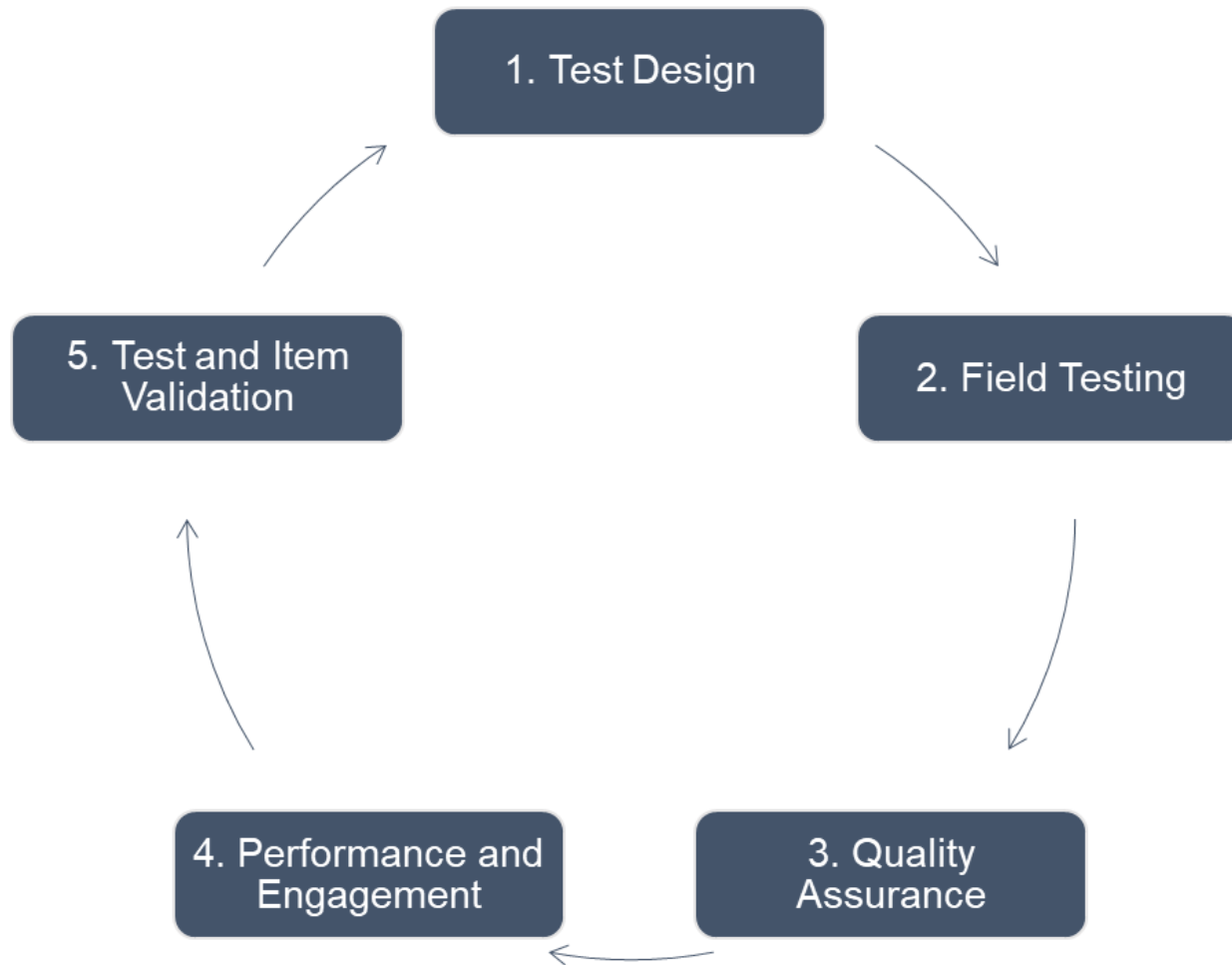
R: In that one.

Slovenia, PIAAC

Questions – On the Uses of Process Data

- What might the uses/applications of process data be in educational assessment?
- What practical, pedagogic and ethical issues might their use raise?

Please identify and explain some uses, and some examples from your own experience.



From Maddox, 'The uses of process data in large-scale educational assessments, 2023 (forthcoming).

References

Bruckner, S. & Pellegrino, J. (2016). 'Integrating the Analysis of Mental Operations Into Multilevel Models to Validate an Assessment of Higher Education Students Competency in Business and Economics'. *Journal of Educational Measurement*. 53 (3), 293–312.

Clark, A. (2011). *Supersizing the Mind: Embodiment, Action and Cognitive Extension*. Oxford University Press.

Embretson, S. (2016). 'Understanding Examinees' Responses to Items: Implications for Measurement'. *Educational Measurement: Issues and Practice*, Vol. 35, No. 3, pp. 6–22.

Ercikan, K., Guo, H., & He, Q. (2020). 'Use of Response Process Data to Inform Group Comparisons and Fairness Research', *Educational Assessment*, 25 (3), 179-197.

Ercikan, K. & Pellegrino, J. 'Validation of Score Meaning using Examinee Response Processes for the Next Generation of Assessments'. In K. Ercikan, and J. Pellegrino (Eds). *Validation of Score Meaning for the Next Generation of Assessments*. Routledge. 1-8.

Ryans, D. J., & Frederiksen, N. (1951). 'Performance tests of educational achievement'. In E. F. Lindquist (Ed.), *Educational measurement* (1st ed., pp. 455–494). Washington, DC: American Council on Education.

Goodwin, C. (2007). 'Environmentally Coupled Gestures', in S. Duncan, J. Cassell, & E.T. Levy (Eds) *Gesture and the Dynamic Dimension of Language*. John Benjamins. 195-212.

Kane, M. & Mislevy, R. (2017). 'Validating Score Interpretations Based on Response Processes'. In K. Ercikan, and J. Pellegrino (Eds). *Validation of Score Meaning for the Next Generation of Assessments*. Routledge. 11-24.

Maddox, B. (2015). 'The neglected situation: assessment performance and interaction in context', *Assessment in Education: Principles, Policy & Practice*, 22:4, 427-443,

Maddox B. (2017). 'Talk and Gesture as Process Data'. *Measurement: Interdisciplinary Research and Perspectives*. 15 (3-4). 113-127.

From Maddox (2023, forthcoming). 'The uses of process data in large-scale educational assessments, OECD working paper. Paris.

Maddox, B. & Zumbo, B.D. (2017). 'Observing Testing Situations: Validation as Jazz' in B.D. Zumbo, and A.M. Hubley (Eds) *Understanding and Investigating Response Processes in Validation Research*. Springer. 179-192.

Newton, P.E. (2019) 'What is response process validation evidence and how important is it? An essay reviewing Ercikan and Pellegrino (2017) and Zumbo and Hubley (2017)', *Assessment in Education: Principles, Policy & Practice*, 26 (2), 245-253.

Oranje et al. (2017) In K. Ercikan, and J. Pellegrino (Eds). *Validation of Score Meaning for the Next Generation of Assessments*. Routledge.

Zumbo, B.D. (2017). 'Trending Away From Routine Procedures, Towards an Ecologically Informed 'In Vivo' View of Validation Practices'. *Measurement: Interdisciplinary Research and Perspectives*, 15:3-4, 137-139.

Zumbo, B.D., Liu, Y., Wu, A.D., Shear, B.R., Astivia, O.L.O. & Ark, T.K. (2015). A Methodology for Zumbo's Third Generation DIF Analyses and the Ecology of Item Responding. *Language Assessment Quarterly*, 12, 136-151.

Zumbo, B.D., Maddox, B., & Care, N. (forthcoming) 'Process and Product in Computer-Based Assessments: Clearing the Ground for a Holistic Validity Framework. *European Journal of Psychological Assessment*'.