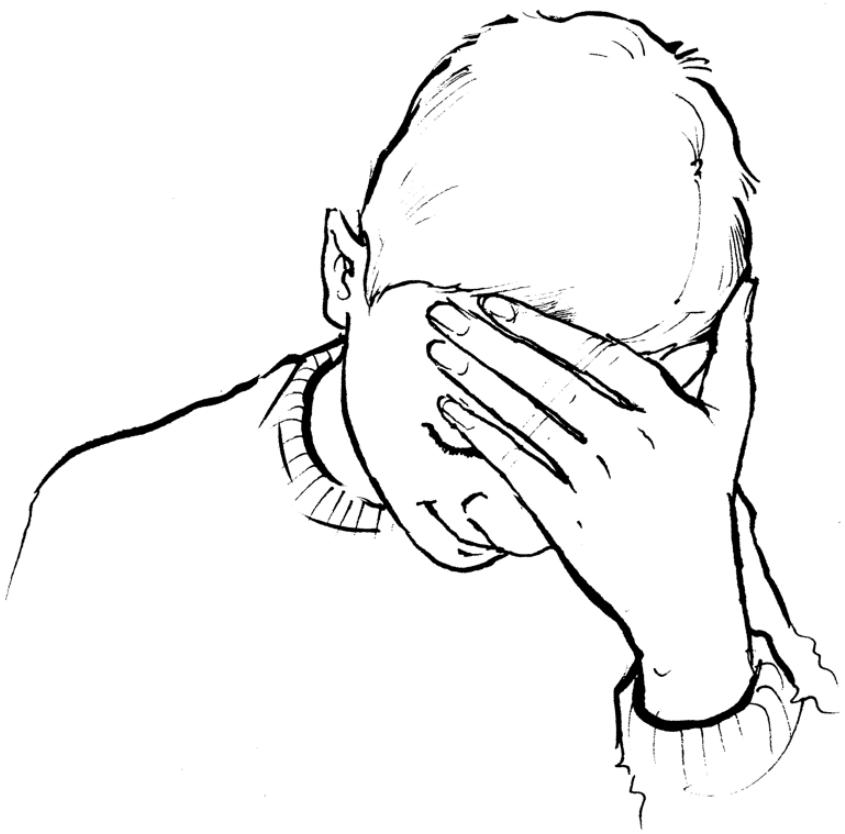
Sensors and Traces: Investigating Assessment Response Processes with Digital Data

Professor Bryan Maddox
Assessment MicroAnalytics & University of East Anglia

Beyond Results 2021: From Log Data to Valid Inferences – Theory Based Construction of Process Indicators



Part 1: *Definitions*

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*.’

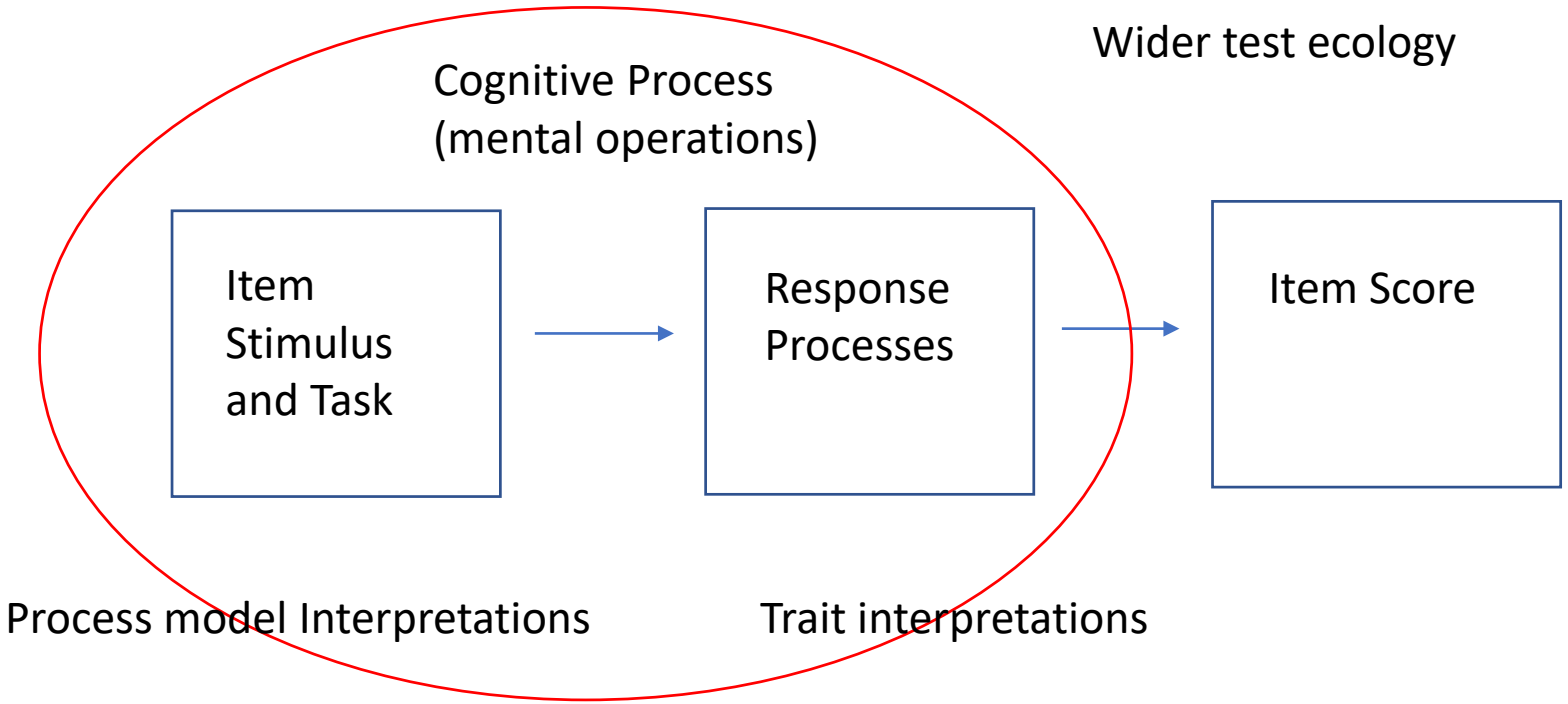
(Ercikan and Pellegrino, 2017, p2).

Defining Response Processes (2)

‘.. 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. This definition expands response processes beyond the cognitive realm to include emotions, motivations and behaviors.’

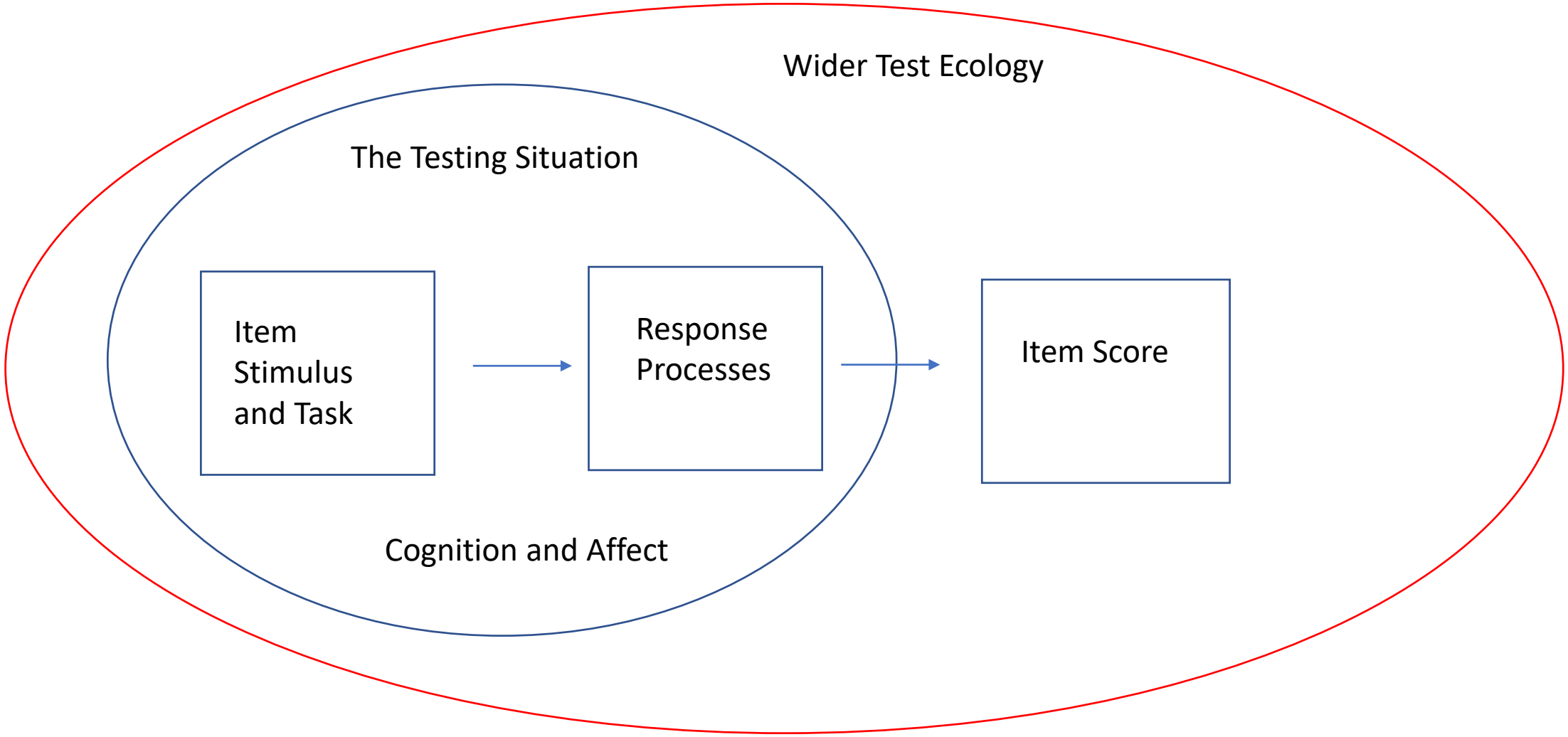
(Hubley & Zumbo, 2017, p. 2)

Cognitive Perspectives



See Embretson (2016); Kane and Mislevy (2017).

Ecological Perspectives



See Zumbo (2017).


MALCHIN

Observing Testing Situations

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.

Test Administrator: *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)

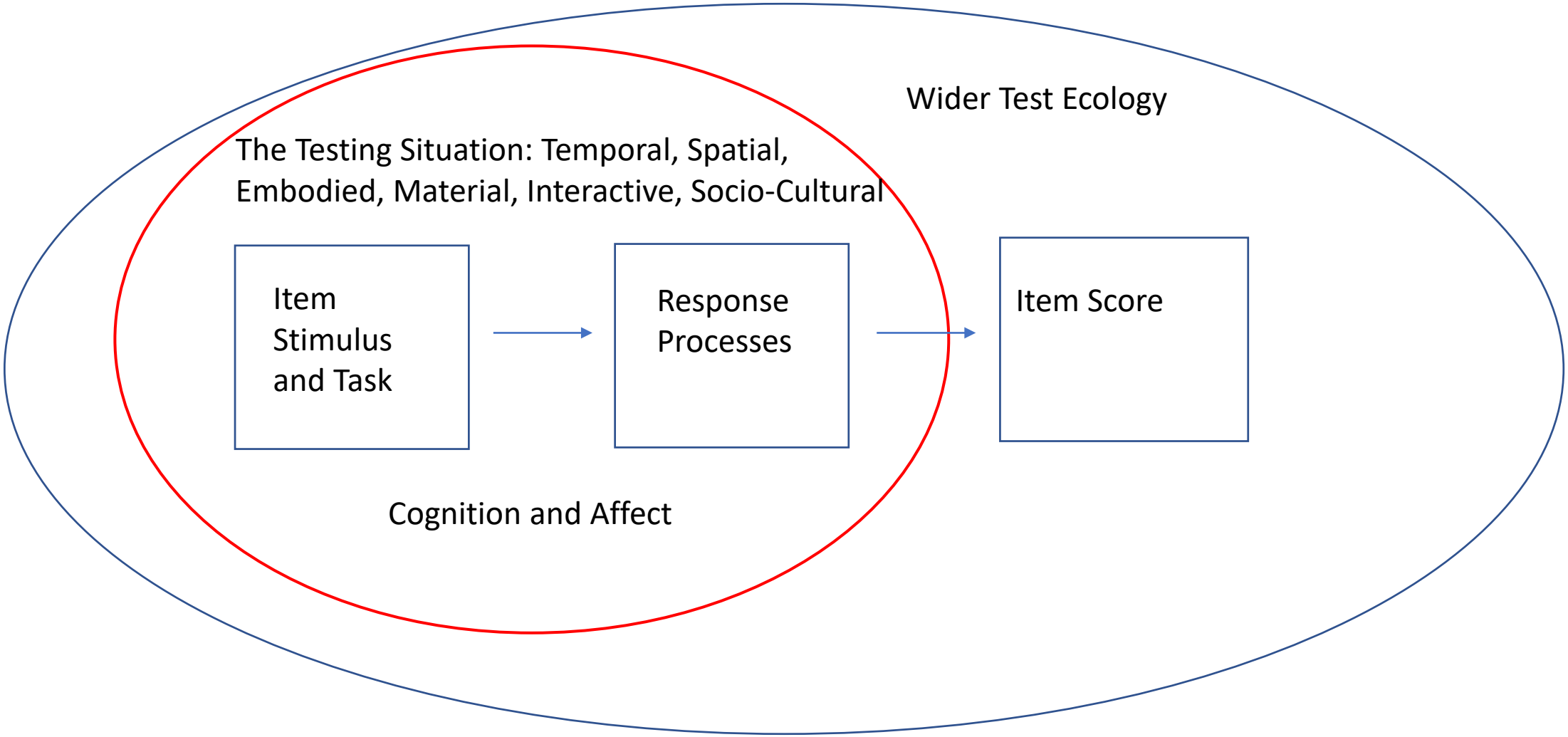
Test Administrator : *If you think it is difficult you can skip the question*. (The respondent is really concentrating on the task. Three herders are now sitting watching the assessment event. A fourth herder enters)

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

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

From Maddox, 2015, p437.

Extended & Situated Cognition



Extended, Embodied, Action-Oriented and Environmentally Coupled Cognition (see Clarke 2011; Goodwin, 2000).

Refining the Definition

Assessment response processes involve the cognitive and affective dimensions of test taking, with respect to the various temporal, spatial, embodied, material, interactive and socio-cultural characteristics of the testing situation. All response process data therefore requires interpretation and validation to establish whether it can be used as an adequate and valid explanation for variation in test scores.

(Maddox, Beyond Results, 2021)

Part 2: *Interpretations*

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).

‘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.’ (Hubley and Zumbo, 2017, p2).

Observed Response Processes

Observed response behaviours are not the ‘traces’ of response processes. They *are* the response processes. The question is not about their presence, but whether their interpretation and use is valid, ethical and fair.

(Maddox, Beyond Results, 2021)

Part 3: *Specifications*

Sensors and Probes

‘…the most illuminating of all […] are direct probes and modelling of the processes underlying test responses [...]. At the simplest level, this might involve querying respondents about their solution processes or asking them to think aloud while responding to exercises during field trials’

(Mesick, 1989. p743)

Micro-Analysis	Meso-Analysis	Macro-Analysis
Micro-analysis involves a sharp focus, and high granularity to capture variation in within item processes.	Medium focus analysis to inform item level judgements and comparisons across items, administrators, sub-groups, and contexts.	Wide angle analysis of whole group and social responses to, and perceptions of the test, its content, uses and consequences.
Typically includes *observational methods* such as eye tracking and log data on keystroke and clickstream data on sequences of action in individual items.	Typically includes *retrospective methods* such as retrospective think aloud, student and administrator questionnaires, and increasingly log data on item response times.	Methods include retrospective surveys, social media analysis, and interviews on user and stakeholder, reception, and public opinion.
High spatial and temporal resolution can capture within item navigation and engagement, scan path sequences and affective response to specific item features.	Temporal and spatial resolution sufficient to inform judgements about test taker understanding and perceptions, and to inform DIF analysis.	Temporal resolution extends across the assessment cycle. Spatial resolution includes institutional practises and perceptions, whole country analysis, and comparisons between countries.

The Spatial and Temporal Resolution of Process Data. (Maddox. Beyond Results, 2021)

Diagram informed by the 'Micro-Macro Validation Continuum' in Newton (2016) 'Micro- and Macro Validation: Beyond the 5 Sources Framework for Classifying Validation Evidence and Analysis.

Conclusions

- Observed behaviours *are* response processes. Our task is to make valid interpretations and uses. i.e., just as we do with test scores.

- Like test scores, the intended uses of process data require validation arguments and evidence.

- We must specify the particular characteristics of process data as we make arguments about its validity, ethics, and fairness.

References

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

Goodwin, C. (2000). 'Action and Embodiment within Situated Human Interaction', Journal of Pragmatics, 32 (10), 1489-1522.

Embretson, S. E. (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.

Ercikan, K. & Pellegrino, J. W. (2017). 'Validation of Score Meaning Using Examinee Response Processes for the Next Generation of Assessments', in Ercikan, K. & Pellegrio, J.W. (Eds). Validation of Score Meaning for the Next Generation of Assessments'. Routledge. New York and London. pp 1-8.

Hubley, A.M. & Zumbo, B.D. (2017). 'Response Processes in the Context of Validity: Setting the Stage'. In Zumbo, B.D. & Hubley, A. M, (Eds). Understanding an Investigating Response Processes in Validation Research. Springer. pp 1-12.

Kane, M. & Mislevy, R. (2017). 'Validating Score Interpretations Based on Response Processes', in in Ercikan, K. & Pellegrio, J.W. (Eds). Validation of Score Meaning for the Next Generation of Assessments'. Routledge. New York and London. pp 10-24.

Newton, P.E. (2016). 'Micro- and Macro Validation: Beyond the 5 Sources Framework for Classifying Validation Evidence and Analysis.

Maddox, B. (2015). 'The Neglected Situation: Assessment Performance and Interaction in Context'. Assessment in Education: Principles, Policy and Practice. Vol. 22 (4) 427–443.

Messick, S. (1989). 'Validity' in Lin, R.L. (Ed.) Educational Measurement. American Council on Education and Macillan, New York.

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.