

Coordination and Validity in Measurement across Science and Medicine: Historical and Epistemological Perspectives

Schedule:

14:30 – 14:45: Welcome and introduction (Michele Luchetti)

14:45 – 15:30: Flavia Padovani (Drexel University) – “From Coordination to Empirical Validity, and Back”

15:30 – 16:15: Leah McClimans (University of South Carolina) – “Measurement, Coordination and Hermeneutics”

16:15 – 16:30: Break

16:30 – 17:15: Cristian Larroulet Philippi and Miguel Ohnesorge (University of Cambridge) – “Is Physical Measurement Relevantly Similar to Human Science Measurement?”

17:15 – 18:00: Uljana Feest (University of Hannover) – “Coordination, Context-Sensitivity, and the Validity of Inferences in Psychology: Beyond Attributes and Instruments”

18:00 – 18:30: General discussion (chair: Michele Luchetti) and informal reception

Abstracts:

Flavia Padovani:

TBA.

Leah McClimans:

The coordination problem asks how we imbue our measuring instruments with empirical significance. In other words, how do we coordinate our measuring instruments with the phenomena we want them to assess? In the empirical literature on measurement, the coordination problem is sometimes discussed in terms of validity, i.e. ensuring a measuring instrument measures what it intends to measure. The problem associated with coordination (or validity) is that it confronts a circle: If I want to know if my measuring instrument does a good job of capturing the phenomena of interest--say temperature or humidity or quality of life--then it seems that I need to know already a great deal about temperature, humidity or quality of life. I need to know, for instance, how temperature fluctuates across locations or people at a single point in time, or how quality of life changes with disease trajectory. Yet this information is precisely what the

measuring instrument is designed to provide. So, how can we ever coordinate our instruments?

To answer this question, I examine Hasok Chang's discussion of coherentism in measurement. As I will illustrate, his proposal has much in common with philosophical hermeneutics (Gadamer, 2004), nonetheless, it emphasizes the stabilization of the hermeneutic circle over time. We might think of this stabilization as a point in time when we know enough about the phenomena of interest such that all the questions we want to ask (for a particular purpose) are answered by the measuring instrument. Once we reach stabilization, if the measuring instrument gives us an answer we don't expect, we tend to call it error or bias. Achieving stability usually means that the phenomena of interest can be standardized, and at least for some metrologists, measurement has been achieved. Yet when we look closer, as I do in this essay, standards get revised, some phenomena are never standardized, some measures are never stabilized, and questions of coordination continue to haunt measurement well-beyond their sell-by date. What is going on? I suggest that the quintessence of measurement is not standardization, but rather hermeneutic dialogue. Sometimes this dialogue becomes stagnant, stability and standardization ensue. But this is the exception and not the rule. Indeed, scientific progress relies on it.

Cristian Larroulet Philippi and Miguel Ohnesorge:

Is Physical Measurement Relevantly Similar to Human Science Measurement?

Disputes about the possibility of quantitative measurement in the human sciences remain. Both optimists and pessimists justify their conclusions by highlighting (dis)similarities to measurement in physics. The debate has been driven by evolving views about which such similarities must obtain for quantification to occur. Lately, pessimists have stressed disanalogies regarding the experimental control over measurands and confounders between psychology and experimental physics (Trendler 2009; 2019).

We discuss the case of seismometry to illustrate a case of successful theory-mediated measurement without experimental control. Looking at this case is worth beyond refuting Trendler—seismometry resembles human measurement in several relevant aspects that are absent in the frequently discussed case of temperature: multidimensional indicators, value-laden considerations, causal complexity, etc. We also discuss the relationship between validity and coordination, and use the seismology case to illustrate what the validity literature fails to illuminate: the semantic openness of the measurand's theoretical definition.

References:

- Trendler, Günter. 2009. 'Measurement Theory, Psychology and the Revolution That Cannot Happen'. *Theory & Psychology* 19 (5): 579–99. <https://doi.org/10.1177/0959354309341926>.
- . 2019a. 'Conjoint Measurement Undone'. *Theory & Psychology* 29: 100–128. <https://doi.org/10.1177/0959354318788729>.

Uljana Feest

Coordination, Context-Sensitivity, and the Validity of Inferences in Psychology: Beyond Attributes and Instruments

In a recent article, Kellen et al (2022) point to the difficulties of establishing that two psychological measurement tools measure the same attribute. The problem, as they construe it, is that (a) even if two instruments measure the same attribute, they may have different “coordination functions,” and (b) even if they have the same coordination function, it is exceedingly difficult to determine what this function is for any given attribute/instrument. In the classic literature about construct validity (e.g., Campbell & Fiske 1959), one approach to this problem is to try and separate out the variance that is due to the attribute from variance that is due to the instrument. However, this focus on attribute and instrument fails to confront yet another problem, namely the potential context-sensitivity of both the instrument and the attribute: Regarding the context-specificity of an attribute, this can take different forms: an attribute might only be triggered in specific contexts, or it can be expressed differently in different contexts, thus requiring different instruments depending on context (Wajnerman-Paz & Rojas-Líbano 2022). In turn, an instrument can be context-sensitive either in the sense of either being particularly reliable for specific contexts or by being particularly prone to confounders in specific contexts.

In my talk, I will argue that the context-sensitivity of both psychological objects and instruments has not been sufficiently recognized, adding another layer of difficulty to the nomic problem (Chang 2004). Building on previous work about the reactivity of the psychological subject matter (Feest 2022), I will argue that in psychological and psychometric research the role of context needs to be explored rather than eliminated. This is because, ultimately, we want theories and instruments that are informative of the real world. I will propose an analysis of experimental inference in psychology that might elucidate how psychologists can address the problem of context-sensitivity.

References:

- Campbell, Donald & Fiske, Donald (1959), Convergent and Discriminant Validation of Traits by the Multitrait-Multimethod Matrix. *Psychological Bulletin* 56(2), 81-105.
- Chang, Hasok (2004), *Inventing Temperature: Measurement and Scientific Progress*. Oxford University Press.
- Feest, Uljana (2022), Data Quality, Experimental Artifacts, and the Reactivity of the Psychological Subject Matter. *European Journal for the Philosophy of Science* 12, 13. <https://doi.org/10.1007/s13194-021-00443-9>
- Kellen, David; Davis-Stober, Clinton P.; Dunn, John; Kalish, Michael (2022), The Problem of Coordination and the Pursuit of Structural Constraints in Psychology. *Perspectives on Psychological Science* 1-12, DOI: 10.1177/174569162097477
- Wajnerman-Paz, Abel; Rojas-Líbano, Daniel (2022), On the role of contextual factors in cognitive neuroscience experiments: a mechanistic approach. *Synthese* (2022) 200:402. <https://doi.org/10.1007/s11229-022-03870-0>