Mensuration without Representation
or: Answering Brandom's Question:
Pragmatism, Measurement, and the Human Sciences

G. A. Reynolds

January 16, 2014

Abstract

Brandom's Question is (roughly): "what features must be exhibited by practices in order that those practices count as having conceptual content?" In the case of measurement, the question may be restated as "what features must be exhibited by our measurement practices so that they be counted as genuinely having the character of measurement?" [TODO: refine this]

Alternatively, Price's Question: something like "what is the science of us that accounts for our conceptual activities?"

thus: "The Pragmatic Question", the question pragmatism asks that integrates philosophy and science; what story can we tell about ourselves that is at the same time a (the right) story about natural science? (See Talisse's intro)

Contents

1 Introduction **4**

2 Paradigmatic Examples **4**

3 Mathematics **5**
- 3.1 An Unnatural History of Number . . . 5
- 3.2 From Number to Structure . . . 6
- 3.3 Mathematical Foundations . . . 6

4 The Ethnomethodological Perspective **7**

5 The Four Scales **8**
- 5.1 Nominal Scales . . . 8
- 5.2 From Ratios to Intervals . . . 9

6 Pragmatism and Measurement **12**

7 Causality and the Space of Reasons **12**

8 Measurement **13**
8.1 Previous Work . . . 14
8.2 Model Theory . . . 15
8.3 Measurement as assignment of numbers . . . 15
8.4 Validity as assessment of correctness . . . 16
8.5 Variables . . . 16
8.6 Error . . . 17

9 Validity, Reliability, Error **18**

10 RCT and Self-validation **19**

11 Vocabularies **19**

12 Conflation of Causal and Logical Relations **19**

13 Deflationism about Validity **19**

14 Fact-Value **20**

15 Word-World **21**

16 Hypothetical Entities **21**

17 Personal v. Subpersonal **22**

18 Spaces **22**
18.1 Natural space of causes . . . 22
18.2 Discursive space of reasons . . . 22

19 Notes **23**
19.1 Evolution . . . 23
19.2 Verum Factum . . . 23
19.3 Rational Evidence . . . 24
19.4 Misc . . . 24
19.5 Erotetic Discursive Practice . . . 25
19.6 Replication . . . 25
19.7 Myths and Mythologies . . . 25
19.8 Dopes . . . 25

20 Mensuration without Representation **26**
20.1 Previous Work . . . 27
20.2 Model Theory . . . 28
20.3 Measurement as assignment of numbers . . . 28
20.4 Validity as assessment of correctness . . . 29
20.5 Variables . . . 30
20.6 Error . . . 30

Appendices **31**

A Bibliography **31**

List of Figures

1 Introduction

First Principle of Measurement: to measure is to compare.

Second Principle of Measurement: to compare is to categorize.

Remark 1 *Measurement essentially involves two measurands. You cannot measure an individuum without involving a second individuum.*

Remark 2 *These principles are not as simple as they look. They involve a fundamental philosophical issue, namely the nature of representation and its relation to practice.*

2 Paradigmatic Examples

Remark 3 *Discussions of measurement are often highly abstract, and depending on the discipline may introduce specialized vocabulary familiar only to readers with experience in the discipline. To make our text intelligible to non-specialists we need specific, detailed, and to the extent possible simple examples of actual measurement.*

- Primitive measurement: counting, height, weight
- Time: undoubtedly the most important development in the history of scientific measurement is the measurement of time. Yet papers on measurement rarely (in my experience, at least) discuss this topic.
- Temperature
- Hardness (Mohs scale)
- Psychology
  - Psychophysics - measurement of perception, e.g. loudness, sweetness, color perception, etc.
  - Personality measurement - anxiety, etc.; the Big Five Personality Traits
  - Intelligence
  - Test theory?
  - Cognitive Psychology?
- Sociology
  - Unemployment rate
  - Religiosity
  - What are the best (simplest, most common, etc.) examples of sociological measures?

- Anthropology
  - Do anthropologists measure? Insofar as they analyze structure, it is in principle possible to see what they do as mapping cultural constructs to mathematical structures; see below.

3 Mathematics

3.1 An Unnatural History of Number

Abstract

Most accounts of measurement, at least in the behavioral sciences, involve some notion of assigning numbers, but usually do not look very closely at the concept of number. Since the term "number" covers a variety of distinct concepts (both historically and theoretically) we should start by getting clear about just what we're talking about.

It's hard to see how to talk about number without circularity, since we do not have an antecedent notion of number that we can apply to all the relevant historical concepts. The history of concepts recognizably involving something like "number" (in the West) is complex so take what follows very schematically. See Ancient Philosophy and Mathematics (MIT Open Courseware); Høyrup, *Lengths, widths, surfaces: a portrait of old Babylonian algebra and its kin*; Grattan-Guinness, "Numbers, Magnitudes, Ratios, and Proportions in Euclid'sElements"; Euclid's Elements. For example, The Greek Conception of Number, gives examples showing that "from the Greeks through at least the middle ages 'one' was not held to be a number".

1. Euclidean "quantities" - three distinct kinds of quantity
   - Arithmetic Number (counting discrete quantities). cf. ἀριθμός (*arithmos*, number). Euclid: "A number is a multitude composed of units."
   - Geometric Magnitude (measuring continuous quantities). Length of a line, area of a square, volume of a solid.
   - Ratio. Harmonic? "[T]he trio of Euclidean quantities, number–magnitude–ratio, surely bears an intentional cultural correlation with three subjects of the Aristotelian quadrivium, arithmetica–geometria–harmonia." Grattan-Guinness, "Numbers, Magnitudes, Ratios, and Proportions in Euclid'sElements", p. 367
2. Zero becomes a number
3. Negative numbers treated as genuine numbers
4. The 19th Century
   (a) Irrationals: in the 19th century Dedekind, Cantor et al. They find a precise definition of irrational numbers as "cuts"; they become bona fide "numbers"

(b) Complex numbers: also in the 19th century, $\sqrt{-1}$ becomes a genuine number

(c) Set theory: quantitative concept of number replaced by structural concept

5. 20th Century

(a) Category theory: alternative foundational theory replaces concept of set with concept of morphism; mathematics as essentially involving structural isomorphisms

(b) Many other classes of mathematical objects and relations; number-qua-quantity (magnitude) discarded

Remark 4 *Operations? Once we move to an algebraic perspective operations are as fundamental as objects. The help define what it is to be an object of a certain type.*

Remark 5 *Counting: some writers confuse categorization and counting. E.g. "According to Mann (2001), a discrete variable assumes values that are obtained from counting, for example, number of houses in a certain block while continuous variables are obtained by measuring and thus, assumes any value contained in an interval, for example, the height of a person. On the other hand, ordinal variables are obtained by ranking. Therefore, discrete and continuous variables are quantitative whereas ordinal variables are qualitative." Yusoff and Mohd Janor, "Generation of an Interval Metric Scale to Measure Attitude", pp. 3-4*

3.2 From Number to Structure

Remark 6 *The historical shift (mainly 19th c.) in mathematics from a focus on quantity and magnitude to a focus on structure and function.*

Significance to measurement: shift in perspective from measurement as quantitative estimation to measurement as structural-functional modeling.

Remark 7 *Compare role of group theory in physics. Algebraic structures treated as models of aspects of nature. This is not reducible to the traditional notion of measurement as the estimation of quantities, but it is no less essential to the science of physics. The lesson is that we should view quantitative estimation (measurement) as a species of the more general notion of structural (algebraic) modeling of nature. After all, measurement scales are themselves algebraic structures (e.g. the real numbers are a field.)*

3.3 Mathematical Foundations

Remark 8 *Why mathematical foundations? Because they afford distinct conceptual orientations. If we adopt an inferential semantics, this means they have different implications. The question is then what practical effect this may (or may not) have for the way we conceptualize measurment, especially in the human or behavioral sciences.*

- Set Theory

- Category Theory
- Type Theory

4 The Ethnomethodological Perspective

Remark 9 *EM studies of mathematics, counting, etc. provide a very different, practice-based (or pragmatist) perspective on measurement.*

5 The Four Scales

Abstract

abstract

Stevens' claim is that the four scales form a hierarchy, wherein each scale subsumes those that precede it. Thus the ratio scale "contains" the features of the nominal, ordinal, and interval scales. True enough, on Stevens' way of looking; but that is not the only way of looking. The argument here is a genealogical perspective is better.

5.1 Nominal Scales

First problem: "nominal" is a misnomer. Nominalism is a philosophical doctrine which claims that ... But when we categorize things, we try to say something about things in the world. It's not just a matter of names.

Second problem: the sense-reference distinction. Names always have both; but for science it is critical that terms be treated as purely extensional. "Electron" is treated as a term whose extension includes electrons and only electrons, full stop; any intensional sense to the term is (in principle) irrelevant to scientific usage. But it works, since we have a pretty good scientific idea of what electrons are.

This works just fine for the hard sciences, but when it comes to the human sciences it is a major problem. What is the extension of, say, "anxiety"? The problem is that it is not even clear that it has an extension. So trying to measure anxiety seems kind of pointless.

Other examples: gender terms; legal status terms (e.g. undocumented, legal permanent resident, etc.) A term like "male" seems to have a pretty clear intensional sense, but it is no small issue to decide on its extension. What about an individual with female sex chromosomes who has male characteristics and functions as a male in society?

The problem goes beyond the trivial observation that terms may have different meanings for different persons. If that were the only issue we could just settle on the correct meaning and classify people who don't agree with it as in error. But the problem is that there is no way to settle on one true correct extensional meaning for any term. Nor is there any reliable way to determine (a priori, as it were) whether a particular use of a term is extensional or intensional. Words inevitably have intensional senses. A person might use "The Fab Four" to refer to John, Paul, George, and Ringo without knowing that they are the Beatles. The classic example is the morning star and the evening star: both refer to the same planet (same extension), but we have no (a priori) way of knowing whether somebody who uses either term is aware of that fact.

A strict formal account of such terms would have to resort to intensional semantics, which is a relatively arcane area of modern logic, involving so-called possible worlds and the like. Not something the average psychologist or sociologist is likely to have mastered.

5.2 From Ratios to Intervals

As a purely historical matter, the so-called ratio scale preceded the interval scale (at least as those scales are conceived by Stevens). The difference is not the presence or absence of "true zero", but the very concept of *zero*. For the Greeks, number *is* quantity, and that is why they did not have a zero. If number is length, then absence of length cannot be a number. The number line has an origin, but that is a very different notion. So the Greek conception of number was essentially the ratio scale, without the concept of zero *as a number*.

Remark 10 *NB: the Kelvin scale has a zero, but it is not possible for anything to actually have a temperature of* 0 deg*K. So Kelvin's zero functions just like the origin of the Greek number line.*

By the same token, our use of "zero" to label a location on an interval scale deceives us. It would be more accurate to call it "center".

Furthermore, note that the Greek conception of a number as a *ratio* represents a radical expansion of the concept of number. Before that, we may imagine, number amounted to quantity. There is a fundamental conceptual (and real?) difference between a discrete quantum (thing), on the one hand, and the ratio on one magnitude to another.

So in terms of Brandom's Question, one thing we must be able to do in order to count as deploying ratios-as-numbers is to compare two distinct individuals under one description. For example, we must be able to 1) treat two distinct rods as being "the same", as both having extent in space, and 2) comparing one to the other (thus ordering), and finally 3) treating one as a unit and the other as a multiple of the unit. All of this was possible even in the absence of an explicitly articulated concept of ratios as numbers; the final essentially creative move involved a change in *discursive* practice: the move to calling ratios numbers. This change in practice *instituted* a change in concepts, rather than the other way around. Creatures incapable of practices involving *treating* distinct things as "the same" (in the appropriate way), comparing them, treating one as a reference unit, etc. would not be capable of coming up with the notion of ratio as number.

So historically the move to ratios-as-numbers involved changes in practices and therefore concepts.

Historically, again, the Arabic/Indic invention of zero cannot be viewed as merely adding a number to the Greek number line. Rather, it involved a fundamental change in the very concept of number. Or more accurately, it involved the emergence of a fundamentally different concept of number, one that accomodated (what we call) rational numbers without the concept of ratio-as-number. The concept of ratio-as-number is conspcuously absent in the earliest book on algebra, al-Khawarizmi's *Kitab al-Jabr wa-l-Muqabala* (from which we have the word "algebra"). What the Greeks conceived of as ratios, the Arabs conceived of as fractions, literally fracturing of a whole, rather than ratios of wholes. The task of algebra is to restore wholes; hence the title of al-Khawarismi's manual: "Concise Book on al-Jabr (literally, bone-setting, mending of fractures) and al-Muqabala (balance)".

There is no way to conceive of zero as a quantity or magnitude, so to characterize precisely the concept of number involved we must characterize the practices involving number in which zero played a role. Here the fundamental innovation seems to involve algebraic calculation, which requires (at least implicitly) negative numbers.

The detailed story of the practical and conceptual innovations associated with the invention of algebra by Arabic-speaking cultural actors is beyond the scope of this paper, but a few brief remarks are in order. Al-Khawarizmi's *Kitab*, the earliest document exhibiting genuinely algebraic manipulations, was a *practical* manual. It contains no abstract mathematical theorizing; where it refers to numbers, it always uses concrete numbers, i.e. numbers *of* things, and the problems it exhibits are mostly matters of commercial or other financial accounting (e.g. dividing inheritances). It does not suggest a notion of zero nor of negative numbers; but, critically, it does involve addition and subtraction of quantities in order to *balance* accounts.

To cut a long story short: the claim here is that the emergence of the abstract mathematical concepts of zero and negative numbers *as numbers* was historically and conceptually parasitic on practices involving addition and subtraction, deficits and surpluses, and the balancing of accounts.

On this view, zero emerges as the center of a structure, flanked on either side by negative and positive numbers. And again, when this notion found its way into the European tradition it eventually generated a fundamental change in the very concept of number - a change that is historically and conceptually parasitic on practices of calculation.

Remark 11 *We have two sorts of genealogy: the historical record of the emergence of practices and concepts, and the structural "genealogy" of the practices that institute conceptual content. By treating conceptual structure as a genealogical structure I mean that we can identify the structure of the concepts involved so as to show how the "later" structures build on the "earlier" ones. This sort of "genealogy" is conceptual and logical (that is, philosophical) rather than historical; what Huw Price calls "philosophical anthropology". But it still counts as a genealogy, in that it tells a story of how we come to have the concepts we have.*

Remark 12 *TODO: show more clearly (in Brandomian terms) how the practical/conceptual structure of interval scales depends on the practical/conceptual structure of ratio scales. This would involve showing how the features of the practices that institute the one concept presuppose those that institute the other. Or something like that.*

Significance for the four scales? For one thing, it reverses the order of ratio and interval scales. The ratio scale is more primitive than the interval scale. It also suggests a change in terminology. The fundamental difference between ratio and interval scales is not ratios v. intervals, but a change in number concept: a ratio scale has an origin (misnamed "true zero"), and an interval scale has a center. A ratio scale measures magnitudes; an interval scale measures directed magnitudes (vectors). So a better nomenclature would use "scalar" and "vector" for "ratio" and "interval", respectively.

Alternatively, an interval scale is essentally an algebraic structure. So we could call ratio scales arithmetic, and interval scales algebraic.

Remark 13 *What about scales that use the complex numbers? Why not?*

Relevance for measurement? Hmmm. At this point I'm not sure if any substantial implications flow from this analysis. But it is a different interpretation of the scales so it could have pedogogical implications.

One possibility: we take interval scales to be about structures with centers (and hence measurements with orientations), rather than about mere intervals. This is a conceptual change. If we suspect some "construct" has interval structure, this means something more than merely that levels are expressed as intervals, or that only certain operations on the scale are allowed. It means that the structure involved has a center, that individuals can have "positive" and "negative" levels of the measurand, and that the additive structure of the scale implicates "balancing" and not merely accumulation.

Remark 14 *"Additive" structure is standardly taken to be an essential property of any empirically measurable attribute. But the very term suggests the arithmetic scale, where addition means accumulation. But the "additivity" of an algebraic scale with a center and "negative measurements" involves moving in both directions, so to speak; not merely accumulation and diminution, but also "negation". Arithmetic addition as a combining operation distinct from the corresponding algebraic operation. The Greeks could not subtract 3 from 2 (this qua characteristic of arithmetic additivity). The Arabs, with an accounting-based, algebraic conception of computation, could, because the result could be conceptualized as a deficit rather than a length.*

Example: use of an interval scale of some sort with questions like "Do you approve/disapprove of the job the President is doing?" On the mathematical interpretation of an interval scale as an algebraic structure with a center, this forces us to treat disapproval as negative approval. That seems dubious; we could just as well treat approval and disapproval as distinct qualities, each involving an arithmetic "ratio" scale with an origin rather than a center.

6 Pragmatism and Measurement

In the case of the scales, we can take Brandom's Question as a schema and ask four distinct questions:

What features must our measuring practices exhibit in order to count as having the significance of:

- Categorization? (Nominal scale)
- Ordering? (Ordinal scale)
- Measuring distance? (Interval scale)
- Measuring ratio? (Ratio scale)

Stevens does not ask Brandom's Question; what matters for him is the mathematical structure of the scales. For example, interval scales do not have a "true" zero, but ratio scales do; thus, Stevens sees a ratio scale as an interval scale augmented by a true zero.

One problem with this mathematicized perspective is that it clashes with history.

But it also has conceptual and logical problems.

We can ask "Is Stevens' concept true (valid, etc.)?" But that seems the wrong question; it is a simple fact that we can describe the four scales in just the way Stevens does. By definition, interval scales do not have a true zero and ratio scales do. That's a purely mathematical matter, and it is virtually always possible to offer alternative descriptions of mathematical structures, no one of which can be selected as the one true description. But the point of the four scales is to address issues of empirical measurement, so the better question is whether the Stevens description is more enlightening or useful than alternatives. The argument here is that it is not, that a pragmatic, genealogical description (Price: philosophical anthropology) offers a better account of measurement. Better in the sense that it focuses on our practices of measurement rather than on the derived abstract description of the mathematical properties of the scales we devise.

7 Causality and the Space of Reasons

Abstract

abstract

Abell, "Narrative Explanation"
Crane and Brewer, "Mental Causation"
Gross, "A Pragmatist Theory of Social Mechanisms"
Jackson, "Mental Causation"
Lowe, "The Causal Autonomy of the Mental"
Lowe, "Non-Cartesian Substance Dualism and the Problem of Mental Causation"
C. MacDonald and G. MacDonald, "Mental Causes and Explanation of Action"
Menzies and Price, "Causation as a Secondary Quality"
Morris, "Causes of Behaviour"
Williamson, "The Broadness of the Mental"

8 Measurement

Abstract

abstract

Remark 15 *Micro-macro: temp as macro v. motion of molecules*

Emergence: liquidity is an emergent property of H2O molecules; is temp an emergent property of moving molecules? It must be insofar as temp is a subjective property (hot, cold, etc.)

Supervenience: or is temperature something that supervenes on groups of molecules in motion?

To measure is to characterize under a mathematical description. Instead of "measurement", use the broader notion of mathematical description. So-called nominal measurement is not quantitative (nor is ordinal measurement); calling it measurement clashes with our intuition, which connects measurement with quantity or magnitude. But both do involve mathematical structure. Mathematics is the science of structure, not quantity.

Measurement claims are thus construed as claims about the structure of some state of affairs in the world. We express such claims in the vocabulary of mathematics (plus an empirical vocabulary involving a "dimension" such as length); a "valid" measurement is a claim expressing or describing a mathematical structure that corresponds accurately (correctly) to the way things are.

Observable v. unobservable: implicit causal relationship. Observable as proxy for unobservable. They must covary.

But this distinction is not simple. Temperature *sensation* is observable, but sensation is distinct from the property in the world. When we measure temperature, we use proxy properties, such as the height of a column of mercury. So temperature is not observable in the required sense. That is, its mathematical structure is not directly observable. Contrast with measurement of length, which is directly observable. Or is it? To measure length we rely on the sensations involved in vision: we see that the measurand is twice the length of the unit instrument. But not really: we do not *see* length per se; rather, we see a stick and use the term "length" to express something about it, based on our experience with things in the world, namely one of the ways we can compare them. Which suggests that terms like "length" are expressive in Brandom's sense: they allow us to say what we can only otherwise do. What we do is compare things; saying that a stick is 1 meter long just saves us the trouble of carrying out a comparison.

Alternatively we could express the same idea in terms of affordances: when we look at a stick, we do not see its length, but we do see (so to speak) one of its affordances: sticks afford lengthwise comparison. (cf. Gibson)

Furthermore, there is the problem of the Myth of the Given and need to explain how we go from merely responding to understanding. This too tends to subvert the observed/unobserved distinction, since we have to ask just what it is that is observed, and what it is to observe. We cannot rely on mere sensory input, since that leads to the Myth. Insofar as observation is a move in the Game of reasons, it is already "theoretical", that is, conceptual, from the very start.

IOW, the observable/unobservable distinction is often conflated with the Given/theoretical distinction. Observables are no more given than unobservables are. But they are directly connected to the

causal order. So it would be better to talk of the distinction between causal and rational orders instead of observables and unobservables. Or perhaps we should stick to vocabulary talk, and make a distinction between observation reports and other sorts of expressions. Some things afford observation reports, others do not.

Electrons are not observables; they do not afford observation reports. But they are causally related to things that do afford such reports. The job of theory is to articulate the hypothesized structure that accounts for such reports in terms of causal relations with electrons. This involves two of the three sorts of language moves: language entries (things affording observation reports), and language-language (theory). Language exits involve what we do, not theoretical predictions about what things in the world do, so the theory predicts future language-entry moves (observations).

This is quite different from e.g. defining SES in terms of occupation, etc. Such definitions are conceptual and do not involve causal relations. Occupation does not cause SES; it is involved in what SES *means* (inferentially), rather than what it is or how it came to be. So defining it is not not about discovering the nature of something in the world. Contrast definition of electron: it must answer to the way things are in the causal order. Our notion of SES must only answer to the way things are in the normative order, which is our order, our way of doing things, the way we cope. If it's useful, we use it; if not, we try other definitions. There is no question of its truth or accurate representation of something in the world. Its a piece of methodological pragmatism: its only purpose is to explain our doings. No metaphysics here, and also no (genuine) measurement. Putative measurements of SES should be treated as methodological conveniences, not as claims about the true state of affairs in the world. Claims that may help us cope or decide what to do, or even predict what will happen. Not because we've measured some fact in the causal order, but because we know something about norms, and norms have a kind of predictive power. We know what ought to be the case; whether things in fact will turn out that way is a different matter.

SES measures as descriptions, which do not necessarily entail predictions. Compare studies primate sociality.

Evolution, selective pressures, etc. Primate anthropologists want to discover selection pressures, not "causes" or the ordinary type. That is causality in evolution is different than causality in physics. Evolutionary causality v. nomological causality. SES measures as a way of getting at "selection pressures" that result in social change, etc.

"We can use the kinds of methods described here to test hypotheses about the selective forces that shape behavioral strategies and to construct comparisons across individuals, groups, or taxa." (Silk et al. p. 223)

8.1 Previous Work

"Paraphrasing N.R. Campbell (Final Report, p.340), we may say that measurement, in the broadest sense, is defined as the assignment of numerals to objects and events according to rules." (Stevens, "On the Theory of Scales of Measurement", p. 677)

"[M]eaningful measurement is possible only if enough is known about the attribute so as to justify its logical operationalization into prescriptions from which a measurement instrument can be devel-

oped." (Sijtsma, "Psychological measurement between physics and statistics", p. 787)

I would rather say the measurement is possible only if we have a theory of description that allows us to make predictions involving measurable (observable) phenomena.

8.2 Model Theory

Truth and consequences and measurement claims.

Relevance of MT: (valid) measurement is all about representation, reference, truth, and validity. (Although a pragmatist might argue it is about what works rather than what corresponds to reality.) Tarski's semantic theory of truth and model-theoretic account of consequence together form the pinacle of this approach.

Tarski (Convention T and model theory) as the pinacle of representational accounts of truth and consequences.

Relevance to measurement? We want to know if our measurement claims are truth, and if the inferences we make involving such claims are valid.

Measurement claims reduce to mathematical claims plus empirical claims. The mathematical part of this accounts for structure.

Model theory: to prove a logical consequence relation between a set of statements Γ and a statement A, first translate them from the formal calculus to the language of ordinary theory (e.g. Group Theory), and then prove the resulting theorems using the informal techniques of the ordinary theory.

Is something similar involved in "proving" an empirical measurement claim (which is a theory)? One difference is goals: the goal of MT is to show that the formal calculus is "good". Science isn't too worried about formal calculi, but it would presumably be a good thing if we could express scientific theories formally and thereby enable formal (automated) reasoning about them. But we don't normally express measurement claims in a formal calculus. Indeed, since measurement claims necessarily involve an empirical component (e.g. units of measure involving empirical properties, that is properties of things in the world), to do so would require formalizing such empirical notions, thus draining them of their empirical content).

8.3 Measurement as assignment of numbers

"Paraphrasing N.R. Campbell (Final Report, p.340), we may say that measurement, in the broadest sense, is defined as the assignment of numerals to objects and events according to rules." (Stevens, 1946, p.677).

This can't be entirely correct. What we assign is not a numeral but a location or position in a mathematical structure. E.g. to assign '3' to a quantity is not to attach a free-standing "numeral" to it, but to assign it a place in the structure of integers.

So each scale type corresponds to a class of mathematical structures.

Nominal: sets? But sets are partially ordered.

Ordinal: sets? But sets also give us intervals?

A nominal scale seems to involve set membership (characteristic functions) at least. But if we can measure the size (cardinality) of a set we end up with order and intervals. So it looks like we must stipulate that these mathematical properties are not to be ascribed to the measurands. Thus nominal measurement involves a partial mapping to sets, or rather a mapping to a set structure that does not admit of ordering or intervals. Hmmm.

Ordinal scales involve order without difference. Again that makes it hard to think of ordinal measurement as involving mapping to sets. Lattice theory?

Does it make sense to think of a mapping to a logical rather than a mathematical structure?

Better: we take set theory a little bit at a time. Start with the basic axioms, then define preorders, posets, etc. So we can treat something as a poset without introducing cardinal and ordinal numbers (I think).

In any case, the upshot is that (representational) measurement postulates a mathematical structure to the measurand.

Michell's concern with whether or not a variable or construct is in fact quantitative can be restated in structural terms. Quantitative properties etc. (in the world) have mathematical structure. Or, to say that something is measurable is to say that it has a particular kind of structure.

Validity "how well the measured variable represents the attribute being measured" comes out as *referential fidelity*. Measurement of something that lacks the requisite mathematical structure will then lack referential fidelity. Referential fidelity is broad enough to cover both accuracy and precision of measurement.

8.4 Validity as assessment of correctness

I.e. to assess something as correct or incorrect is to measure it against a norm. In the case of e.g. temperature measurement, the norm is the "true" temperature of the sample being measured.

Relevance: validity involves normativity and a kind of measurement against (usually unstated) norms or "true" standards, which may be (idealized) methods, etc.

Thus referential fidelity as correctness of representation.

References:

H. Chang, *Inventing temperature*

H. Chang, "Measurement, Justification, and Scientific Progress"

H. Chang, "Spirit, Air, and Quicksilver"

A. Martin and Lynch, "Counting Things and People"

Michell, "Normal Science, Pathological Science and Psychometrics"

Sherry, "Thermoscopes, thermometers, and the foundations of measurement"

See British Journal of Psychology, Aug 1997 vol 88 issue 3: Michell, "Quantitative science and the definition of measurement in psychology" and six commentaries.

8.5 Variables

References:

Schwarz, "Is Psychology Based on a Methodological Error?"
Toomela, "Variables in Psychology"
Stam, "The Fault is Not in Ourselves, but in Our Methods"

8.6 Error

References:
T. W. Smith, "Refining the Total Survey Error Perspective"

9 Validity, Reliability, Error

Remark 16 *What is the point of worrying about validity? Is it something in the world that we are trying to discover? Then we're trying to find "the right description of the world" (Putnam). Or is it a concept, so that validity talk is about conceptual analysis and definition?*

Or: we try to find the right description, and validity talk is part of how we decide that we have found it.

Remark 17 *Why do psychometricians and the like worry so about validity?*

Hypothesis: when they say "validity", what they're really interested in is scientific legitimacy. Effectively, to say that a test (etc.) is valid is to say that it is in fact scientific. Thats the practical import of the concept of validity for them.

Unpack this. Expose the assumptions and implications.

key concepts:

- validity treated as a special kind of property - of what?
- constructs
- (latent) variables
- indicators

"validity" as code for:

- legitimacy
- vindication
- credibility
- proof (good premises + valid inference)

Remark 18 *On the idea that validity something (a property, etc.) that we look for in scientific theories in order to distinguish good ones from bad: see Putnam on fact/value distinction. We use value judgments - simplicity, parsimony, etc. - in every aspect of science (thought), esp. in weeding out bad theories. For there is no external or objective criterion of acceptability for theories to which we can appeal, nor is there any such citerion that does not involve value judgments.*

Remark 19 *So along with the fact/value distinction, and the analytic/synthetic distinction, the internal/external distinction also collapses? Or do we just exclude the notion of external? No; we need to retain the idea of an external world that is independent of us and to which some of our judgments are answerable. We don't get to just make stuff up and call it true (correct) for at least*

some of our claims. There is no external absolute authority that can decide for us which theories are true, or rather which we should endorse, but that does not mean there is no external world that is authoritative for some of our sayings. But isn't that trying to have it both ways? How can our theories answer to the world if we cannot appeal to the world or some other external authority to sort them out? See Brandom.

Related issue: what counts as evidence? How do we decide? What are we doing when we decide that something counts as strong (weak) evidence in support of a theory? What are the criteria of adequacy for an account of evidence?

10 RCT and Self-validation

See Cartwright on RCT as self-validating. This seems to mean that RCTs are valid by construction.

This nicely parallels industrial QA notions of guaranteeing quality by designing a production process that prevents defects.

What's the logic here? Is self-validation really possible? How can a process validate itself - isn't the very idea inherently circular? Or rather, don't we land in a regress? After all, if the idea is to specify a process that yields validity, how do we know that that process is itself valid?

11 Vocabularies

Measurement as description. Description v. evaluation. Price on naturalisms. The bifurcation thesis.

12 Conflation of Causal and Logical Relations

13 Deflationism about Validity

Remark 20 *Deflationism seems to depend essentially on some form of expressivism. Or maybe they amount to the same thing?*

How can we get out of this mess? One way is to deflate the notion of validity, just deny that it is a substantive property. When we claim that a result is valid etc. what we are really saying is that we endorse it, approve of it, etc. It's an expressive device. Compare the semantic deflationist's idea that calling something true amounts to endorsing or approving of it.

So if we discard the notion of validity (since it does no real work), don't we find ourselves lacking something essential? Well, we just need a vocabulary that allows us to say explicitly the sorts of things we find it useful to be able to express with respect to a study or qx technique. For example: credibility, utility, legitimacy, vindication, justification, etc.

Remark 21 *The notion of validity seems to be connected to the problem of deciding which theories we should endorse. What are the criteria of adequacy for any notion (or theory) of validity?*

Or: what are the requirements that should be met by any purported explanation of validity? Both particular cases and the general idea. Tarski gives us something like this for logical validity; what about "validity" as the term is used by psychometricians, test theorists, etc.?

Contrast: claims of validity for a case, v. explanation of what validity is.

The objection will no doubt be that we need some kind of standard, which is just to say that we want to measure this something (validity, credibility, whatever). Implicit in all this is the notion that there is some "objective" fact of the matter to which our study/technique/etc. is ansswerable. A study is valid iff - what? If it meets some definite "objective" criteria. Methodological criteria, conditions of validity, etc. In the psychometrics and testing tradition this appeal to external authority is expressed as something along the lines of "measures what it purports to measure". Which is only meaningful insofar as a) there is actually something there to measure, and b) it is in fact susceptibel to measurement.

And usually this is expressed in statistical terms. But that dog won't hunt either - you cannot get to validity via statistics. All you can do is measure central tendencies and variance - not enough to establish validity, which is a substantive notion. (analysis elsewhere).

To say that sth is valid is just to say that it is admirable (Peirce?), or perhaps that it is virtuous, that it has the virtues we prize.

14 Fact-Value

Messick, for one, conflates two kinds of fact/value distinction. The Kantian idea that we structure our own experience (etc.), Sellars' Myth of the Given, and etc. - such stuff shows how there is no data that is "objective" and given i.e. "data is theory-laden".

So facts involve what Putnam calls "epistemic values".

Messick confuses epistemic and ethical values. He seems to think that although we cannot arrive at value-free facts, this is because brute facts are always packaged with ethical values. The idea seems to be that ethical values are something separate from facts but always attached to them somehow. Whereas the real problem is that there is no genuine distinction between fact and (epistemic) value. Facts express (as it were) our epistemic values.

Messick's confusion is clear in his distinction between the scientific and social "roles" of validity - as if the social (value-laden) aspect of (Messickian) validity is something distinct from the science. "[I]t is fundamental that score validation is an empirical evaluation of the meaning and consequences of measurement. As such, validation combines scientific inquiry with rational argument to justify (or nullify) score interpretation and use." (p. 742) But "scientific inquiry" and "rational argument" are not two distinct things that can be combined. They are the same thing, at least conceptually. If there is a difference here, it is sociological - science as a way of conduction oneself, etc.

Messickian validity boils down to some notion of empirical support for theoretical explanations. For him "evidential basis" seems to correspond to "real" science, and "consequential basis" to "rational argument".

"[B]oth meaning and values are integral to the concept of validity..." (p. 747). The problem here is that the contrast with value is fact, not meaning.

"Meaning" is not something that can be empirically "validated".

15 Word-World

One problem with e.g. Messick is fuzziness about the relation of language to world. Ditto for any notion of "measuring a concept".

Re: validity: is it supposed to be a property of things in the world, or just a concept? Per Messick, validity is "associated with" score interpretation and use. This would seem to imply that it is a matter of language (concepts). But the language is just sloppy; "score interpretation" might (should) refer to how we take a score to relate to some fact in the world, in which case the question is just what is validity-in-the-world.

In any case, Messick's whole discussion is muddled on this point; it is rarely clear when he is talking about facts, concepts, or the relation between the two. Is a "construct" supposed to be something in the world or a concept the describes some aspect of the world?

Construct v. "indicators".

Compare positivist notions of observational language v. theoretical language. So-called indicators are (I understand) supposed to be empirical observables. Their relation to the construct is (must be) a matter of theory; but then is that theoretical (conceptual) structure to be taken as a mirror of reality, such that the construct is a real (albeit "hidden") bit of the world and its relations to the indicators are real relations in the world?

16 Hypothetical Entities

Putnam, Brandom, etc. - if the existence of (some) hypotheticals makes no difference in the way things are then we can just discard them. As Putnam puts it, "Would mathematics *work* one bit less well if these funny objects *stopped* existing? Those who posit "abstract entities" to account for the success of mathematics do not claim that we (or any other things in the empirical world) *interact* with the abstract entities. But if any entities do not interact with us or with the empirical world *at all*, then doesn't it follow that *everything wouuld be the same if they didn't exist*?" (Collapse, p. 33)

This points out another problem with e.g. latent variables, namely that they are supposed to have causal powers, but, insofar as they are abstract at least, they have no connection to the empirical world and so cannot cause anything. The counterargument would presumably be that hidden does not necessarily mean abstract. But in that case they must have a location in space-time, even if we don't know what it is. But this just leads to more problems: where are hidden psychological processes supposed to occur? It can't be the brain, since they are (by stipulation) psychological, not neurological.

So it seems we have no choice but to treat postulation of hidden stuff as a matter of Brandomian methodological pragmatism: useful, but without ontological consequences. "Constructs" may be useful for explaining observable indicators, but they don't really exist in any meaningful sense. But the usual story goes the other way around: indicators are useful because they are how we get constructs.

Another perspective: hard science starts with observation and moves to number, theory, etc. Psychometrics reverses this, starting with number and theory (latent vars, etc.) and then seeking observational support.

Example: temperature v. anxiety. The former is directly associated with publicly available bodily experience. Is the latter? Anxiety may be experienced by individuals but it is not public in the way the temperature of an external phenomenon is public. The sensation of temperature may be private, but it is directly linked to the (public) causal order. So although both are essentially conceptual, only the former answers to the state of the world. There is no prima facie reason to think that the concept of anxiety represents something in the world; in this respect it is just like "Zeus" or "phlogiston". So trying to measure "it" inescapably involves starting with a speculative ontological hypothesis. Whereas trying to measure temperature starts with something observable.

Same point made from perspective of anthropology: we can be 100% confident that all cultures encounter things in the world that are hot or cold, regardless of their concepts. But we cannot be sure that anxiety - either the thing itself or the concept/term - is a cultural universal.

Remark 22 *TODO: explicit comparison of psychometrics with failed but arguably scientific measurement projects like the caloric theory, phlogiston, etc. on the one hand, and clearly pseudo-scientific projects like astrology, ESP, etc. on the other. The task is to determine which one psychometrics is. Is psychometrics in a "caloric theory" phase, genuinely scientific yet lacking good theories, or is it like astrology? It's open to the psychometrician to argue that the science is young, and that just because it is a science it will self-correct, so that eventually we will have the theories and practices needed to make precise measurements of anxiety (or its successor concept) - possibly by measuring brain states and structures. The problem with this argument is that it continues to overlook or ignore the fact that categories like "anxiety" are only intelligible in the space of reasons; they are creatures of the normative order, not of the natural (causal) order. So the question becomes whether they answer to anything in the causal order in the way that temperature answers to physical states of the world. The fundamental hypothesis (or speculation) of psychometrics seems to be that the concepts of the folk psychology from which psychometrics draws its "constructs" are causally related in some way to the causal order. The objection is that there is no such causal relation, that the relations between these concepts is entirely normative.*

17 Personal v. Subpersonal

Reasons v. causes

18 Spaces

18.1 Natural space of causes

18.2 Discursive space of reasons

19 Notes

19.1 Evolution

Instead of ”the QA process”, the proper object of investigation is the local evolution of discourse.

EM studies local produced order. It may come up with a structural description. But locally produced order is the outcome of an essentially evolutionary process - the mutual adaptation of the participants to each other and the context. Also, any such model may not (probably will not) generalize. But what does generalize is the evolutionary mechanism itself, just like in biology.

Rational selection as the mechanism of the evolution of discursive performances. What accounts for the deontic attitudes we adopt regarding performances? Brandom’s account describes the architecture of such posturings and the significances the institute. But it does not really address the logic of discourse as an evolutionary process.

The idea is that Brandom provides an account of discourse qua rational action. Different attitudes are endorsed or undertaken for reasons - that is the source or ground of the intelligibility of discursive practice. So if we view the unfolding of discourse as being governed by the logic of evolution, we can treat Brandom’s sort of rational pragmatism as the selection mechanism that accounts for why some attitudes (meanings) survive (are endorsed) and others do not. Meanings that survive must fit into the space of reasons - they must be assertable and justifiable, even if the participants are unable to explicitly articulate this. This makes the evolution of discourse intelligible as a rational process, rather than a natural process. Responses to questions are not explicable as effects caused by ”true values” or the like; this would make them fundamentally non-rational. Or to borrow a bon mot from Garfinkel, this would make respondents ”rational dopes”.

Similar language: ”negotiation”, e.g. ”...I suggest that the content of talk indicates that imposed hierarchies are continually re-negotiated...” Negotiation as rational evolution?

The ”true score” and other orthodox models account for sentience, not sapience.

19.2 Verum Factum

Cartesianism (spectator, etc.) inspection, discovery, certainty, foundationism (external foundation grounding knowledge) v.

Verum Factum, geneological/historical, following growth/development, not certainty but ???; no foundationism, no priviledged vocab, no external source of authority

Critical notions: authority. For evidence etc. key idea is authority - the only kind of authority is the kind we assent to. So the question is what do we treat as authoritative and why, rather than how can we discover the One True external foundational source of authority and learn to speak its language

Critical notions: vocabulary. Regardless of what there is, we can only talk about it by using vocabs.

Relevance to SR: we make our truths, by engaging in dialog with respondents in order to teach/train them to understand what we want. In other words we work to make our scorecards converge. We can never be sure that researchers and respondents understand each other, have the same interpretations of qx text, etc. But we can do what nature does in evolution and learning: institute a cyclic process of

experiment, feedback, and correction. This is operational even at the most simple and basic level of communication. So we can use this fact to our advantage.

Communication interactions as not essentially different from processes of evolution and learning. Evolutionary process tend to coordinate organism and environment; learning processes adapt the learner to the task environment, etc. Any discursive exchange - even simple greetings, etc. - does the same sort of thing: coordinate and mutually adjust the parties to the exchange.

19.3 Rational Evidence

Evidence-Based Rational SR

RCT: isolate the causal factor that links Treatment to Outcome

THe mistake make by orthodox SR (shown by its vocab of measurement, error, etc.) is that it confuses the space of causes and the space of reasons.

In RCT, we observe a stimulus followed by a response (T followed by O) and postulate a causal relation. In SR, we observe a Q performance followed by a R performance. In fact this is an idealization since Q and R cannot be isolated - they are both joint performances. Ignore that for now; the point is that what makes them intelligible as performances is the space of reasons, not causes. That is, as discursive episodes they are essentially rational in a way the T-O trials are not. By definition, ”rational” means involving concepts. Stimulus-response does not involve concepts and so is not rational in this favored sense. The natural world may be lawful, but it is not rational.

So SR should abandon the orthodox vocab of measurment, etc. in favor of one involving rationality. What would ”evidence-based” mean, then? Not the kind of evidence involve in natural science, since such evidence does not involve concepts and thus meaning. Instead evidence inescapably involves meaning and understanding. What counts as evidence is what we count as a rational explanation or story. And this necessarily involves the perspective of the participants - it is their rationality, their giving and asking for reasons, that provides the observational basis of evidence.

One consequence: Qx does not involve measurement. SR can use stats to statistically measure the collected data, but that is quite separate from whether the data measure anything. So you can say that x% of resondents pick option X, but that does not mean that you have measured the distribution of ”true values” of some latent variable. What you have measure is a distribution of deontic scores, or discursive postures. There is no warrant for claiming that each member of the x% means the same thing by picking X.

19.4 Misc

1. What is a question? Better: what counts as a question, what is it to ask a question?

2. Ditto for answer.

Q and A as parts of a whole (holistic view)

Q token v. Q performance, etc.

19.5 Erotetic Discursive Practice

EDP as production of data rather than discovery of truth

19.6 Replication

Goal is replication. Compare: blood work, e.g. measuring cholesteral. The measuring apparatus reacts to the sample, not the other way around. For EDP, respondent reacts to the question, so the question is analogous to the blood sample. The response is a kind of measurement of the question, not the other way around.

Replicability means same setup, same experimental conditions; in EDP this means replication of conceptual structure, which is accomplished by the dialog preceding the question. Traditionally, ”ask the same question”; in practice this is impossible, since what counts is not the question text but respondent’s grasp of the sense. So the ”experimental setup” should be viewed as the work of teaching the respondent what the sense of the question is. Survey interviewing is essentially interventionist, but this is not necessarily a bad thing, since lab experiments are too - they ”intervene” to set up experimental ”initial conditions”. The difference is that setting up initial conditions (”same meaning”) in question asking means tutoring the respondent.

19.7 Myths and Mythologies

- The Myth of Question Independence says that the meaning of a question is independent of context. But the meaning of a question is always dependent on what came before it.
- Myth of Autonomy. Interviewer and Respondent.
- Myth of Error

19.8 Dopes

Garfinkel’s dopes - cultural, judgmental, psychological

Dehumanization. Orthodox Survey Research (OSR) dehumanizes participants. The R is a sampling unit. The mythology of OSR measurement treats the human R as a natural object to be measured rather than a person.

20 Mensuration without Representation

Remark 23 *Measurement as a species of modeling.*

Quantities, "levels", etc. - too narrow. E.g. gravity as a force, force as a quantity; but Einstein taught us that gravity is not in fact a force. Or: caloric, heat as a fluid, which has quantity. But heat is not a fluid. Moral: quantification is based on analogy.

Micro-macro: temp as macro v. motion of molecules

Emergence: liquidity is an emergent property of H2O molecules; is temp an emergent property of moving molecules? It must be insofar as temp is a subjective property (hot, cold, etc.)

Supervenience: or is temperature something that supervenes on groups of molecules in motion?

To measure is to characterize under a mathematical description. Instead of "measurement", use the broader notion of mathematical description. So-called nominal measurement is not quantitative (nor is ordinal measurement); calling it measurement clashes with our intuition, which connects measurement with quantity or magnitude. But both do involve mathematical structure. Mathematics is the science of structure, not quantity.

Measurement claims are thus construed as claims about the structure of some state of affairs in the world. We express such claims in the vocabulary of mathematics (plus an empirical vocabulary involving a "dimension" such as length); a "valid" measurement is a claim expressing or describing a mathematical structure that corresponds accurately (correctly) to the way things are.

Observable v. unobservable: implicit causal relationship. Observable as proxy for unobservable. They must covary.

But this distinction is not simple. Temperature *sensation is observable, but sensation is distinct from the property in the world. When we measure temperature, we use proxy properties, such as the height of a column of mercury. So temperature is not observable in the required sense. That is, its mathematical structure is not directly observable. Contrast with measurement of length, which is directly observable. Or is it? To measure length we rely on the sensations involved in vision: we see that the measurand is twice the length of the unit instrument. But not really: we do not see length per se; rather, we see a stick and use the term "length" to express something about it, based on our experience with things in the world, namely one of the ways we can compare them. Which suggests that terms like "length" are expressive in Brandom's sense: they allow us to say what we can only otherwise do. What we do is compare things; saying that a stick is 1 meter long just saves us the trouble of carrying out a comparison.*

Alternatively we could express the same idea in terms of affordances: when we look at a stick, we do not see its length, but we do see (so to speak) one of its affordances: sticks afford lengthwise comparison. (cf. Gibson)

Furthermore, there is the problem of the Myth of the Given and need to explain how we go from merely responding to understanding. This too tends to subvert the observed/unobserved distinction, since we have to ask just what it is that is observed, and what it is to observe. We cannot rely on mere sensory input, since that leads to the Myth. Insofar as observation is a move in the Game of reasons, it is already "theoretical", that is, conceptual, from the very start.

IOW, the observable/unobservable distinction is often conflated with the Given/theoretical distinction. Observables are no more given than unobservables are. But they are directly connected to the causal order. So it would be better to talk of the distinction between causal and rational orders instead of observables and unobservables. Or perhaps we should stick to vocabulary talk, and make a distinction between observation reports and other sorts of expressions. Some things afford observation reports, others do not.

Electrons are not observables; they do not afford observation reports. But they are causally related to things that do afford such reports. The job of theory is to articulate the hypothesized structure that accounts for such reports in terms of causal relations with electrons. This involves two of the three sorts of language moves: language entries (things affording observation reports), and language-language (theory). Language exits involve what we do, not theoretical predictions about what things in the world do, so the theory predicts future language-entry moves (observations).

This is quite different from e.g. defining SES in terms of occupation, etc. Such definitions are conceptual and do not involve causal relations. Occupation does not cause SES; it is involved in what SES means (inferentially), rather than what it is or how it came to be. So defining it is not not about discovering the nature of something in the world. Contrast definition of electron: it must answer to the way things are in the causal order. Our notion of SES must only answer to the way things are in the normative order, which is our order, our way of doing things, the way we cope. If it's useful, we use it; if not, we try other definitions. There is no question of its truth or accurate representation of something in the world. Its a piece of methodological pragmatism: its only purpose is to explain our doings. No metaphysics here, and also no (genuine) measurement. Putative measurements of SES should be treated as methodological conveniences, not as claims about the true state of affairs in the world. Claims that may help us cope or decide what to do, or even predict what will happen. Not because we've measured some fact in the causal order, but because we know something about norms, and norms have a kind of predictive power. We know what ought to be the case; whether things in fact will turn out that way is a different matter.

SES measures as descriptions, which do not necessarily entail predictions. Compare studies primate sociality.

Evolution, selective pressures, etc. Primate anthropologists want to discover selection pressures, not "causes" or the ordinary type. That is causality in evolution is different than causality in physics. Evolutionary causality v. nomological causality. SES measures as a way of getting at "selection pressures" that result in social change, etc.

"We can use the kinds of methods described here to test hypotheses about the selective forces that shape behavioral strategies and to construct comparisons across individuals, groups, or taxa." (Silk et al. p. 223)

20.1 Previous Work

"Paraphrasing N.R. Campbell (Final Report, p.340), we may say that measurement, in the broadest sense, is defined as the assignment of numerals to objects and events according to rules." (Stevens, "On the Theory of Scales of Measurement", p. 677)

"[M]eaningful measurement is possible only if enough is known about the attribute so as to justify its logical operationalization into prescriptions from which a measurement instrument can be developed." (Sijtsma, "Psychological measurement between physics and statistics", p. 787)

I would rather say the measurement is possible only if we have a theory of description that allows us to make predictions involving measurable (observable) phenomena.

20.2 Model Theory

Truth and consequences and measurement claims.

Relevance of MT: (valid) measurement is all about representation, reference, truth, and validity. (Although a pragmatist might argue it is about what works rather than what corresponds to reality.) Tarski's semantic theory of truth and model-theoretic account of consequence together form the pinacle of this approach.

Tarski (Convention T and model theory) as the pinacle of representational accounts of truth and consequences.

Relevance to measurement? We want to know if our measurement claims are truth, and if the inferences we make involving such claims are valid.

Measurement claims reduce to mathematical claims plus empirical claims. The mathematical part of this accounts for structure.

Model theory: to prove a logical consequence relation between a set of statements Γ and a statement A, first translate them from the formal calculus to the language of ordinary theory (e.g. Group Theory), and then prove the resulting theorems using the informal techniques of the ordinary theory.

Is something similar involved in "proving" an empirical measurement claim (which is a theory)? One difference is goals: the goal of MT is to show that the formal calculus is "good". Science isn't too worried about formal calculi, but it would presumably be a good thing if we could express scientific theories formally and thereby enable formal (automated) reasoning about them. But we don't normally express measurement claims in a formal calculus. Indeed, since measurement claims necessarily involve an empirical component (e.g. units of measure involving empirical properties, that is properties of things in the world), to do so would require formalizing such empirical notions, thus draining them of their empirical content).

20.3 Measurement as assignment of numbers

"Paraphrasing N.R. Campbell (Final Report, p.340), we may say that measurement, in the broadest sense, is defined as the assignment of numerals to objects and events according to rules." (Stevens, 1946, p.677).

This can't be entirely correct. What we assign is not a numeral but a location or position in a mathematical structure. E.g. to assign '3' to a quantity is not to attach a free-standing "numeral" to it, but to assign it a place in the structure of integers.

So each scale type corresponds to a class of mathematical structures.

Nominal: sets? But sets are partially ordered.

Ordinal: sets? But sets also give us intervals?

A nominal scale seems to involve set membership (characteristic functions) at least. But if we can measure the size (cardinality) of a set we end up with order and intervals. So it looks like we must stipulate that these mathematical properties are not to be ascribed to the measurands. Thus nominal measurement involves a partial mapping to sets, or rather a mapping to a set structure that does not admit of ordering or intervals. Hmmm.

Ordinal scales involve order without difference. Again that makes it hard to think of ordinal measurement as involving mapping to sets. Lattice theory?

Does it make sense to think of a mapping to a logical rather than a mathematical structure?

Better: we take set theory a little bit at a time. Start with the basic axioms, then define preorders, posets, etc. So we can treat something as a poset without introducing cardinal and ordinal numbers (I think).

In any case, the upshot is that (representational) measurement postulates a mathematical structure to the measurand.

Michell's concern with whether or not a variable or construct is in fact quantitative can be restated in structural terms. Quantitative properties etc. (in the world) have mathematical structure. Or, to say that something is measurable is to say that it has a particular kind of structure.

Validity "how well the measured variable represents the attribute being measured" comes out as referential fidelity. Measurement of something that lacks the requisite mathematical structure will then lack referential fidelity. Referential fidelity is broad enough to cover both accuracy and precision of measurement.

20.4 Validity as assessment of correctness

I.e. to assess something as correct or incorrect is to measure it against a norm. In the case of e.g. temperature measurement, the norm is the "true" temperature of the sample being measured.

Relevance: validity involves normativity and a kind of measurement against (usually unstated) norms or "true" standards, which may be (idealized) methods, etc.

Thus referential fidelity as correctness of representation.

References:

H. Chang, Inventing temperature

H. Chang, "Measurement, Justification, and Scientific Progress"

H. Chang, "Spirit, Air, and Quicksilver"

A. Martin and Lynch, "Counting Things and People"

Michell, "Normal Science, Pathological Science and Psychometrics"

Sherry, "Thermoscopes, thermometers, and the foundations of measurement"

See British Journal of Psychology, Aug 1997 vol 88 issue 3: Michell, "Quantitative science and the definition of measurement in psychology" and six commentaries.

20.5 Variables

References:
Schwarz, "Is Psychology Based on a Methodological Error?"
Toomela, "Variables in Psychology"
Stam, "The Fault is Not in Ourselves, but in Our Methods"

20.6 Error

References:
T. W. Smith, "Refining the Total Survey Error Perspective"

Appendices

A Bibliography

[1] Peter Abell. "Narrative Explanation: An Alternative to Variable-Centered Explanation?" In: *Annual Review of Sociology* 30 (Jan. 2004), pp. 287–310. URL: http://www.jstor.org/stable/29737695 (visited on 12/08/2013) (cit. on p. 12).

[2] E. Andreouli. "Identity, Positioning and Self-Other Relations". In: *Papers on Social Representations* 19.1 (2010), pp. 14.1–14.13.

[3] Michael Bacon. *Pragmatism: an introduction*. Cambridge, UK ; Malden, MA: Polity, 2012. ISBN: 9780745646640.

[4] Patrick Baert. *Philosophy of the Social Sciences: Towards Pragmatism*. en. Polity, Nov. 2005. ISBN: 9780745622477.

[5] Patrick Baert. "Pragmatism as a Philosophy of the Social Sciences". en. In: *European Journal of Social Theory* 7.3 (Aug. 2004), pp. 355–369. DOI: 10.1177/1368431004044198. (Visited on 12/24/2012).

[6] Patrick Baert. "Pragmatism, Realism and Hermeneutics". In: *Foundations of Science* 8.1 (Mar. 2003), pp. 89–106. DOI: 10.1023/A:1022445814115. URL: http://link.springer.com.proxy.uchicago.edu/article/10.1023/A:1022445814115 (visited on 12/08/2012).

[7] Patrick Baert. "Realism versus Pragmatism – an Introduction". en. In: *Foundations of Science* 8.1 (Mar. 2003), pp. 1–2. DOI: 10.1023/A:1022491321867. URL: http://link.springer.com.proxy.uchicago.edu/article/10.1023/A:1022491321867 (visited on 01/27/2013).

[8] Patrick Baert. "Towards a Pragmatist-Inspired Philosophy of Social Science". In: *Acta Sociologica* 48.3 (Sept. 2005), pp. 191–203. DOI: 10.2307/20059943. URL: http://www.jstor.org/stable/20059943 (visited on 12/24/2012).

[9] Mikhail Mikhaĭlovich Bakhtin. *Speech genres and other late essays*. Trans. by Vern W. McGee. University of Texas Press, 1986. ISBN: 9780292775602.

[10] Mikhail Mikhaĭlovich Bakhtin. "The Problem of Speech Genres". In: *Speech Genres and Other Late Essays*. University of Texas Press, 1986, pp. 60–102. ISBN: 9780292775602.

[11] Barry Barnes. "Ethnomethodology as Science". In: *Social Studies of Science* 15.4 (Nov. 1985). ArticleType: book-review / Full publication date: Nov., 1985 / Copyright © 1985 Sage Publications, Ltd., pp. 751–762. ISSN: 0306-3127. URL: http://www.jstor.org/stable/285404 (visited on 01/22/2012).

[12] Randy Elliot Bennett. “Consequences That Cannot Be Avoided: A Response to Paul Newton”. In: *Measurement: Interdisciplinary Research & Perspective* 10.1-2 (2012), pp. 30–32. DOI: 10.1080/15366367.2012.686865.

[13] T. J. Berard. “Rethinking Practices and Structures”. In: *Philosophy of the Social Sciences* 35.2 (June 2005), pp. 196–230. ISSN: 0048-3931, 1552-7441. DOI: 10.1177/0048393105275290. URL: http://pos.sagepub.com.proxy.uchicago.edu/content/35/2/196 (visited on 11/15/2012).

[14] Richard J. Bernstein. *The Pragmatic Turn*. Polity, Apr. 2010.

[15] Jack Bilmes. “Ethnomethodology, Culture, and Implicature: Toward an Empirical Pragmatics”. In: *Pragmatics* 3.4 (Feb. 2010). ISSN: 1018-2101. URL: http://elanguage.net/journals/index.php/pragmatics/article/viewArticle/181.

[16] Paul Black. “EPMA Professionals—Servants or Masters?” In: *Measurement: Interdisciplinary Research & Perspective* 10.1-2 (2012), pp. 33–37. DOI: 10.1080/15366367.2012.677342.

[17] Simon Blackburn. “Invited Introduction: Finding Psychology”. In: *The Philosophical Quarterly* 36.143 (Apr. 1986), pp. 111–122. DOI: 10.2307/2219763. URL: http://www.jstor.org/stable/2219763 (visited on 12/08/2013).

[18] Simon Blackburn. “The Steps from Doing to Saying”. In: *Proceedings of the Aristotelian Society (Hardback)* 110.1pt1 (2010), pp. 1–13. DOI: 10.1111/j.1467-9264.2010.00276.x. URL: http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9264.2010.00276.x/abstract (visited on 12/08/2013).

[19] S. F. Blinkhorn. “Past imperfect, future conditional: Fifty years of test theory”. In: *British Journal of Mathematical and Statistical Psychology* 50.2 (1997), pp. 175–185. DOI: 10.1111/j.2044-8317.1997.tb01139.x.

[20] David Bloor. “Wittgenstein and the priority of practice”. In: *The Practice Turn in Contemporary Theory*. Routledge, 2001, pp. 103–114.

[21] Herbert Blumer. “Sociological Analysis and the ”Variable””. In: *American Sociological Review* 21.6 (Dec. 1956), pp. 683–690. DOI: 10.2307/2088418.

[22] Denny Borsboom. “The attack of the psychometricians”. In: *Psychometrika* 71.3 (Sept. 2006), pp. 425–440. DOI: 10.1007/s11336-006-1447-6.

[23] Denny Borsboom. “Whose Consensus Is It Anyway? Scientific Versus Legalistic Conceptions of Validity”. In: *Measurement: Interdisciplinary Research & Perspective* 10.1-2 (2012), pp. 38–41. DOI: 10.1080/15366367.2012.681971.

[24] Denny Borsboom and Gideon J. Mellenbergh. “Why Psychometrics is Not Pathological A Comment on Michell”. In: *Theory & Psychology* 14.1 (Feb. 2004), pp. 105–120. DOI: 10.1177/0959354304040200.

[25] Denny Borsboom, Gideon J. Mellenbergh, and Jaap van Heerden. “The theoretical status of latent variables”. In: *Psychological Review* 110.2 (2003), pp. 203–219. DOI: 10.1037/0033-295X.110.2.203.

[26] Tom Bramley. ““Measurement” and “Construct” Need to Be Clarified First. Commentary on Newton, P. E. “Clarifying the Consensus Definition of Validity””. In: *Measurement: Interdisciplinary Research & Perspective* 10.1-2 (2012), pp. 42–45. DOI: 10.1080/15366367.2012.677344.

[27] Robert Brandom. *Articulating Reasons : An Introduction to Inferentialism*. Cambridge, Mass.; London: Harvard University Press, 2001.

[28] Robert Brandom. *Between saying and doing : towards an analytic pragmatism*. Oxford; New York: Oxford University Press, 2008.

[29] Robert Brandom. *Making it explicit: reasoning, representing, and discursive commitment*. Harvard University Press, June 1998.

[30] Robert Brandom. *Perspectives on Pragmatism : Classical, Recent, and Contemporary*. Cambridge, Mass.: Harvard University Press, 2011.

[31] Robert Brandom. “Précis of Making It Explicit”. In: *Philosophy and Phenomenological Research* 57.1 (Mar. 1997). ArticleType: book-review / Full publication date: Mar., 1997 / Copyright © 1997 International Phenomenological Society, pp. 153–156. DOI: 10.2307/2953784. URL: http://www.jstor.org/stable/2953784 (visited on 12/31/2012).

[32] Robert Brandom. “The Pragmatist Enlightenment (and its Problematic Semantics)”. In: *European Journal of Philosophy* 12.1 (2004), pp. 1–16. DOI: 10.1111/j.0966-8373.2004.00196.x.

[33] Robert Brandom. “The Social Anatomy of Inference”. In: *Philosophy and Phenomenological Research* 53.3 (1993), pp. 661–666. DOI: 10.2307/2108089.

[34] Robert B. Brandom. “Analyzing Pragmatism: Pragmatics and Pragmatisms”. In: *Perspectives on Pragmatism*. Cambridge Mass. ; London: Harvard University Press, 2011.

[35] Robert B. Brandom. “Classical American Pragmatism: The Pragmatist Enlightenment - and Its Problematic Semantics”. In: *Perspectives on Pragmatism*. Cambridge Mass. ; London: Harvard University Press, 2011.

[36] Robert B. Brandom. “Introduction: From German Idealism to American Pragmatism - and Back”. In: *Perspectives on Pragmatism*. Cambridge Mass.; London: Harvard University Press, 2011.

[37] Robert B. Brandom. “Linguistic Pragmatism and Pragmatism about Norms: An Arc of Thought from Rorty’s Eliminative Materialism to his Pragmatism”. In: *Perspectives on Pragmatism*. Cambridge Mass. ; London: Harvard University Press, 2011.

[38] Robert B. Brandom. “Pragmatism, Expressivism, and Anti-Representationalism: Local and Global Possibilities”. In: *Perspectives on Pragmatism*. Cambridge Mass. ; London: Harvard University Press, 2011.

[39] Robert B Brandom. *Reason in philosophy*. Cambridge, Mass. [etc.]: The Belknap Press of Harvard University Press, 2009.

[40] Robert B. Brandom. “Vocabularies of Pragmatism: Synthesizing Naturalism and Historicism”. In: *Perspectives on Pragmatism*. Cambridge Mass. ; London: Harvard University Press, 2011.

[41] Henry Braun. “Conceptions of Validity: The Private and the Public”. In: *Measurement: Interdisciplinary Research & Perspective* 10.1-2 (2012), pp. 46–49. DOI: 10.1080/15366367.2012.679159.

[42] Svend Brinkmann. “Psychology’s Facts and Values: A Perennial Entanglement”. In: *Philosophical Psychology* 18.6 (2005), pp. 749–765. DOI: 10.1080/09515080500355244.

[43] Svend Brinkmann. “The ethical subject: Accountability, authorship, and practical reason”. In: *SATS* 11.1 (Sept. 2010), pp. 75–89. URL: http://www.degruyter.com.proxy.uchicago.edu/view/j/sats.2010.11.issue-1/sats.2010.007/sats.2010.007.xml (visited on 04/21/2012).

[44] Svend Brinkmann. “The Mind as Skills and Dispositions: On Normativity and Mediation”. In: *Integrative Psychological and Behavioral Science* (Aug. 2011). ISSN: 1932-4502, 1936-3567. DOI: 10.1007/s12124-011-9183-6. URL: http://www.springerlink.com.proxy.uchicago.edu/content/g2222336gjn7x61k/ (visited on 01/02/2012).

[45] Svend Brinkmann. “Towards an Expansive Hybrid Psychology: Integrating Theories of the Mediated Mind”. In: *Integrative Psychological and Behavioral Science* 45.1 (Oct. 2010), pp. 1–20. ISSN: 1932-4502, 1936-3567. DOI: 10.1007/s12124-010-9146-3. URL: http://www.springerlink.com.proxy.uchicago.edu/content/ejq3033k4l08g430/ (visited on 02/04/2012).

[46] Mary Bucholtz and Kira Hall. “Identity and interaction: a sociocultural linguistic approach”. In: *Discourse Studies* 7.4-5 (Oct. 2005), pp. 585–614. ISSN: 1461-4456, 1461-7080. DOI: 10.1177/1461445605054407. URL: http://dis.sagepub.com.proxy.uchicago.edu/content/7/4-5/585.short (visited on 02/12/2012).

[47] Graham Button, ed. *Ethnomethodology and the Human Sciences*. Cambridge; New York: Cambridge University Press, 1991. ISBN: 9780521389525.

[48] Edward G. Carmines and Richard A. Zeller. *Reliability and Validity Assessment*. SAGE, Nov. 1979.

[49] Robyn Carston. “Linguistic Communication and the Semantics/Pragmatics Distinction”. In: *Synthese* 165.3 (Dec. 2008). ArticleType: research-article / Full publication date: Dec., 2008 / Copyright © 2008 Springer, pp. 321–345. ISSN: 0039-7857. URL: http://www.jstor.org/stable/40271119.

[50] Miguel Castillo-Díaz and José-Luis Padilla. “How Cognitive Interviewing can Provide Validity Evidence of the Response Processes to Scale Items”. In: *Soc Indic Res* (), pp. 1–13. DOI: 10.1007/s11205-012-0184-8.

[51] A M Chang, J P Chau, and E Holroyd. “Translation of questionnaires and issues of equivalence”. In: *J Adv Nurs* 29.2 (Feb. 1999). PMID: 10197930, pp. 316–322. URL: http://www.ncbi.nlm.nih.gov/pubmed/10197930 (visited on 08/28/2011).

[52] Hasok Chang. *Inventing temperature: measurement and scientific progress*. Oxford studies in philosophy of science. Oxford ; New York: Oxford University Press, 2004. ISBN: 0195171276 (cit. on pp. 16, 29).

[53] Hasok Chang. “Measurement, Justification, and Scientific Progress”. In: *Inventing temperature: measurement and scientific progress*. Oxford studies in philosophy of science. Oxford ; New York: Oxford University Press, 2004, pp. 220–234. ISBN: 0195171276 (cit. on pp. 16, 29).

[54] Hasok Chang. “Spirit, Air, and Quicksilver”. In: *Inventing temperature: measurement and scientific progress*. Oxford studies in philosophy of science. Oxford ; New York: Oxford University Press, 2004, pp. 57–102. ISBN: 0195171276 (cit. on pp. 16, 29).

[55] Jeasik Cho and Allen Trent. “Validity in qualitative research revisited”. In: *Qualitative Research* 6.3 (2006), pp. 319–340. DOI: 10.1177/1468794106065006.

[56] Lindsey Churchill. “Ethnomethodology and Measurement”. In: *Social Forces* 50.2 (Dec. 1971), pp. 182–191. DOI: 10.2307/2576936.

[57] Aaron V. Cicourel. “Interviews, Surveys, and the Problem of Ecological Validity”. In: *The American Sociologist* 17.1 (Feb. 1982), pp. 11–20. URL: http://www.jstor.org/stable/27702491 (visited on 03/11/2012).

[58] Andy Clark. *Being there : putting brain, body, and world together again*. Cambridge, Mass. [u.a.]: MIT Press, 1997. ISBN: 0262032406 9780262032407 0262531569 9780262531566.

[59] Andy Clark and Josefa Toribio. “Doing without Representing?” In: *Synthese* 101.3 (Dec. 1994), pp. 401–431. DOI: 10.1007/BF01063896.

[60] Herbert H. Clark. *Using language*. Cambridge University Press, May 1996. ISBN: 9780521567459.

[61] David A. Cook and Thomas J. Beckman. “Current Concepts in Validity and Reliability for Psychometric Instruments: Theory and Application”. In: *The American Journal of Medicine* 119.2 (Feb. 2006), 166.e7–166.e16. DOI: 10.1016/j.amjmed.2005.10.036.

[62] Deborah J. Coon. “Standardizing the Subject: Experimental Psychologists, Introspection, and the Quest for a Technoscientific Ideal”. In: *Technology and Culture* 34.4 (Oct. 1993), p. 757. DOI: 10.2307/3106414.

[63] Jeff Coulter. “Cognition: cognition in an ethnomethodological mode”. In: *Ethnomethodology and the Human Sciences*. Ed. by Graham Button. Cambridge; New York: Cambridge Univ. Press, 1991, p. 176. ISBN: 9780521389525.

[64] Jeff Coulter. “Discourse and Mind”. In: *Human Studies* 22.2/4 (Oct. 1999), pp. 163–181. URL: http://www.jstor.org/stable/20011238.

[65] Jeff Coulter. “Language Without Mind”. English. In: *Conversation and Cognition*. Ed. by Hedwig te Molder and Jonathan Potter. Cambridge; New York: Cambridge Univ. Press, 2005, pp. 79–92.

[66] Jeff Coulter. “Logic: ethnomethodology and the logic of language”. In: *Ethnomethodology and the Human Sciences*. Ed. by Graham Button. Cambridge; New York: Cambridge Univ. Press, 1991, pp. 20–50. ISBN: 9780521389525.

[67] Jeff Coulter. “What Is ”Discursive Psychology”?” In: *Human Studies* 27.3 (Jan. 2004), pp. 335–340. URL: http://www.jstor.org/stable/20010378.

[68] Angélique O. J. Cramer. “Why the Item “23 +1” Is Not in a Depression Questionnaire: Validity From a Network Perspective”. In: *Measurement: Interdisciplinary Research & Perspective* 10.1-2 (2012), pp. 50–54. DOI: 10.1080/15366367.2012.681973.

[69] Tim Crane and Bill Brewer. “Mental Causation”. In: *Proceedings of the Aristotelian Society, Supplementary Volumes* 69 (Jan. 1995), pp. 211–253. ISSN: 0309-7013. URL: http://www.jstor.org/stable/4107076 (visited on 12/08/2013) (cit. on p. 12).

[70] Lee J. Cronbach and Paul E. Meehl. “Construct validity in psychological tests”. In: *Psychol Bull* 52.4 (July 1955). PMID: 13245896, pp. 281–302. URL: http://www.ncbi.nlm.nih.gov/pubmed/13245896 (visited on 03/17/2012).

[71] Simon Cullen. “Survey-Driven Romanticism”. en. In: *Review of Philosophy and Psychology* 1.2 (June 2010), pp. 275–296. DOI: 10.1007/s13164-009-0016-1.

[72] Bronwyn Davies and Rom Harré. “Positioning: The Discursive Production of Selves”. en. In: *Journal for the Theory of Social Behaviour* 20.1 (Mar. 1990), pp. 43–63. DOI: 10.1111/j.1468-5914.1990.tb00174.x.

[73] G. Scott Davis. *Believing and Acting: The Pragmatic Turn in Comparative Religion and Ethics*. en. Oxford University Press, Mar. 2012.

[74] Robyn M. Dawes. “Psychological measurement”. In: *Psychological Review* 101.2 (1994), pp. 278–281. DOI: 10.1037/0033-295X.101.2.278.

[75] Daniel C. Dennett. “Styles of Mental Representation”. In: *Proceedings of the Aristotelian Society*. New Series 83 (Jan. 1982), pp. 213–226. ISSN: 00667374. URL: http://www.jstor.org.proxy.uchicago.edu/stable/4545000 (visited on 02/19/2011).

[76] Jacques Derrida. “Signature Event Context”. In: *Limited Inc*. Evanston, Ill.: Northwestern Univ. Press, 1988.

[77] Vincent Descombes. *The Mind’s Provisions: a Critique of Cognitivism*. New French thought. Princeton, N.J: Princeton University Press, 2001. ISBN: 0691001316.

[78] Robert F DeVellis. *Scale development : theory and applications*. Thousand Oaks, Calif.: Sage Publications, 2003. ISBN: 0761926046 9780761926047 0761926054 9780761926054.

[79] Santo Di Nuovo. "Variables and Quantity: What Else?" In: *Integrative Psychological and Behavioral Science* 43.1 (Nov. 2008), pp. 84–88. DOI: 10.1007/s12124-008-9081-8.

[80] Morris Dickstein. *The Revival of Pragmatism: New Essays on Social Thought, Law, and Culture*. Duke University Press, Nov. 1998.

[81] Eli Dresner. "Language and the Measure of Mind". In: *Mind & Language* 25.4 (Sept. 2010), pp. 418–439. DOI: 10.1111/j.1468-0017.2010.01396.x.

[82] Hubert L. Dreyfus. "How Representational Cognitivism Failed and is being replaced by Body/World Coupling". In: *After Cognitivism*. Springer Netherlands, 2009, pp. 39–73. URL: http://www.springerlink.com.proxy.uchicago.edu/content/q4kt33m51hh41456/abstract/ (visited on 10/16/2012).

[83] Hubert L. Dreyfus. "Intelligence without representation – Merleau-Ponty's critique of mental representation The relevance of phenomenology to scientific explanation". In: *Phenomenology and the Cognitive Sciences* 1.4 (Dec. 2002), pp. 367–383. DOI: 10.1023/A:1021351606209. URL: http://dx.doi.org/10.1023/A:1021351606209.

[84] Judith A. Droitcour, Eric M. Larsen, and Fritz J. Scheuren. "The Three-Card Method: Estimating Sensitive Survey Items—With Permanent Anonymity of Response". In: *Proceedings of the Annual Meeting of the American Statistical Association*. Aug. 2001. URL: http://www.amstat.org/sections/srms/proceedings/y2001/Proceed/00582.pdf.

[85] Judith A. Droitcour and Eric M. Larson. "An Innovative Technique for Asking Sensitive Questions: the Three-Card Method". In: *Bulletin de Méthodologie Sociologique* 75.1 (July 2002), pp. 5–23. DOI: 10.1177/075910630207500103. URL: http://bms.sagepub.com/content/75/1/5.abstract (visited on 10/25/2011).

[86] William Egginton and William Egginton Mike Sandbothe. *The Pragmatic Turn in Philosophy: Contemporary Engagements Between Analytic and Continental Thought*. SUNY Press, Apr. 2004.

[87] Susan E. Embretson. "Construct Validity: A Universal Validity System or Just Another Test Evaluation Procedure?" In: *Educational Researcher* 36.8 (Nov. 2007), pp. 449–455. DOI: 10.3102/0013189X07311600.

[88] Mustafa Emirbayer and Douglas W. Maynard. "Pragmatism and Ethnomethodology". In: *Qualitative Sociology* 34.1 (Nov. 2010), pp. 221–261. DOI: 10.1007/s11133-010-9183-8.

[89] George Engelhard and Nadia Behizadeh. "Epistemic Iterations and Consensus Definitions of Validity". In: *Measurement: Interdisciplinary Research & Perspective* 10.1-2 (2012), pp. 55–58. DOI: 10.1080/15366367.2012.681974.

[90] Christopher Essex and William E. Smythe. "Between Numbers and Notions A Critique of Psychological Measurement". In: *Theory & Psychology* 9.6 (Dec. 1999), pp. 739–767. DOI: 10.1177/0959354399096002.

[91] Ronald Fischer and Taciano L. Milfont. “Standardization in psychological research.” In: *International Journal of Psychological Research* 3.1 (July 2010), pp. 88–96. URL: http://mvint.usbmed.edu.co:8002/ojs/index.php/web/article/view/463 (visited on 12/09/2013).

[92] J A Flaherty et al. “Developing instruments for cross-cultural psychiatric research”. In: *J. Nerv. Ment. Dis.* 176.5 (May 1988). PMID: 3367140, pp. 257–263. URL: http://www.ncbi.nlm.nih.gov/pubmed/3367140 (visited on 10/23/2011).

[93] J. A. Fodor. “Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology”. In: *Behavioral and Brain Sciences* 3.01 (1980), pp. 63–73. DOI: 10.1017/S0140525X00001771.

[94] Melissa Freeman. “Validity in Dialogic Encounters With Hermeneutic Truths”. In: *Qualitative Inquiry* 17.6 (July 2011), pp. 543–551. DOI: 10.1177/1077800411409887.

[95] Barbara Fultner. “Inferentialism and Communicative Action: Robust Conceptions of Intersubjectivity”. In: *Philosophical Studies* 108.1 (Mar. 2002), pp. 121–131. DOI: 10.1023/A:1015768316114.

[96] Mary J. Gallant and Sherryl Kleinman. “Symbolic Interactionism vs. Ethnomethodology”. In: *Symbolic Interaction* 6.1 (May 1983). ArticleType: research-article / Full publication date: Spring 1983 /, pp. 1–18. ISSN: 0195-6086. URL: http://www.jstor.org/stable/10.1525/si.1983.6.1.1 (visited on 01/14/2012).

[97] U.S. General Accounting Office (GAO). *Estimating the Undocumented Population: A “Grouped Answers” Approach to Surveying Foreign-Bord Respondents*. Tech. rep. GAO-06-775. Washington D.C.: U.S. General Accounting Office, Sept. 2006. URL: http://www.gao.gov/new.items/d06775.pdf.

[98] Harold Garfinkel. *Ethnomethodology’s program : working out Durkeim’s aphorism*. Ed. by Anne Warfield Rawls. Lanham, Md.: Rowman & Littlefield Publishers, 2002. ISBN: 0742516415 9780742516410 0742516423 9780742516427.

[99] Harold Garfinkel. *Studies in ethnomethodology*. Cambridge, UK : Polity Press, 1984.

[100] Charles Goodwin. “The Interactive Construction of a Sentence in Natural Conversation”. In: *Everyday Language: Studies in Ethnomethodology*. Irvington Publishers, 1979, pp. 97–121. ISBN: 9780470266700.

[101] Joanna S. Gorin. “Reconsidering Issues in Validity Theory”. In: *Educational Researcher* 36.8 (Nov. 2007), pp. 456–462. DOI: 10.2307/4621100. URL: http://www.jstor.org/stable/4621100 (visited on 01/27/2013).

[102] Ivor Grattan-Guinness. “Numbers, Magnitudes, Ratios, and Proportions in Euclid’sElements: How Did He Handle Them?” In: *Historia Mathematica* 23.4 (Nov. 1996), pp. 355–375. DOI: 10.1006/hmat.1996.0038 (cit. on p. 5).

[103] H. P. Grice. “Meaning”. In: *The Philosophical Review* 66.3 (July 1957). ArticleType: research-article / Full publication date: Jul., 1957 / Copyright © 1957 Duke University Press, pp. 377–388. ISSN: 00318108. URL: http://www.jstor.org/stable/2182440.

[104] H. P. Grice. “Utterer’s Meaning and Intention”. In: *The Philosophical Review* 78.2 (Apr. 1969). ArticleType: research-article / Full publication date: Apr., 1969 / Copyright © 1969 Duke University Press, pp. 147–177. ISSN: 00318108. URL: http://www.jstor.org/stable/2184179.

[105] H. P. Grice. “Utterer’s Meaning, Sentence-Meaning, and Word-Meaning”. In: *Foundations of Language* 4.3 (Aug. 1968). ArticleType: research-article / Full publication date: Aug., 1968 / Copyright © 1968 Springer, pp. 225–242. ISSN: 0015900X. URL: http://www.jstor.org/stable/25000329.

[106] H. Paul Grice. *Studies in the Way of Words*. Harvard University Press, 1989.

[107] Paul Grice. “Logic and Conversation”. In: *Studies in the Way of Words*. Cambridge Mass ;;London: Harvard University Press, 1989.

[108] Neil Gross. “A Pragmatist Theory of Social Mechanisms”. en. In: *American Sociological Review* 74.3 (June 2009), pp. 358–379. DOI: 10.1177/000312240907400302. (Visited on 10/10/2012) (cit. on p. 12).

[109] Robert M Groves. “Measurement Error Across the Disciplines”. In: *Measurement Errors in Surveys*. John Wiley & Sons, Inc., 1991.

[110] Susan Haack and Robert Edwin Lane. *Pragmatism, old & new: selected writings*. Prometheus Books, Apr. 2006.

[111] Brian D. Haig. “From Construct Validity to Theory Validation”. In: *Measurement: Interdisciplinary Research & Perspective* 10.1-2 (2012), pp. 59–62. DOI: 10.1080/15366367.2012.681975.

[112] J. Haiman. “Ritualization and the Development of Languages”. In: *Perspectives on Grammaticalization*. Amsterdam: John Benjamins, 1994.

[113] John Hale. “Uncertainty about the Rest of the Sentence”. In: *Cognitive Science* 30.4 (2006), pp. 643–672. URL: http://www.eric.ed.gov/ERICWebPortal/detail?accno=EJ747089 (visited on 03/19/2012).

[114] Rom Harré. “Discursive Psychology and the Boundaries of Sense”. In: *Organization Studies* 25.8 (Oct. 2004), pp. 1435–1453. DOI: 10.1177/0170840604046351. URL: http://oss.sagepub.com/content/25/8/1435.abstract (visited on 02/06/2012).

[115] Rom Harré. “The Discursive Creation of Human Psychology”. In: *Symbolic Interaction* 15.4 (Nov. 1992). ArticleType: research-article / Full publication date: Winter 1992 /, pp. 515–527. ISSN: 0195-6086. URL: http://www.jstor.org/stable/10.1525/si.1992.15.4.515 (visited on 01/19/2012).

[116] Rom Harré et al. “Recent Advances in Positioning Theory”. In: *Theory & Psychology* 19.1 (Feb. 2009), pp. 5–31. ISSN: 0959-3543, 1461-7447. DOI: 10.1177/0959354308101417. URL: http://tap.sagepub.com.proxy.uchicago.edu/content/19/1/5 (visited on 01/23/2013).

[117] John Heritage. *Garfinkel and ethnomethodology*. Cambridge [Cambridgeshire]; New York, N.Y.: Polity Press, 1984. ISBN: 0745600603 9780745600604 0745600611 9780745600611.

[118] C. S. Herrman. “Fundamentals of Methodology Part I: Definitions and First Principles”. In: *SSRN eLibrary* (Apr. 2009). URL: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1373976 (visited on 01/21/2012).

[119] S. Brian Hood. “In Defense of an Instrument-Based Approach to Validity”. In: *Measurement: Interdisciplinary Research & Perspective* 10.1-2 (2012), pp. 63–65. ISSN: 1536-6367. DOI: 10.1080/15366367.2012.681976.

[120] S. Brian Hood. “Validity in Psychological Testing and Scientific Realism”. In: *Theory & Psychology* 19.4 (Aug. 2009), pp. 451–473. DOI: 10.1177/0959354309336320.

[121] Lisa Tsoi Hoshmand. “Can Lessons of History and Logical Analysis Ensure Progress in Psychological Science?” In: *Theory & Psychology* 13.1 (Feb. 2003), pp. 39–44. DOI: 10.1177/0959354303131003.

[122] David Houle et al. “Measurement and Meaning in Biology”. In: *The Quarterly Review of Biology* 86.1 (Mar. 2011), pp. 3–34. DOI: 10.1086/658408.

[123] Dorothy Howie and Michael Peters. “Positioning Theory: Vygotsky, Wittgenstein and Social Constructionist Psychology”. In: *Journal for the Theory of Social Behaviour* 26.1 (Mar. 1996), pp. 51–64. ISSN: 1468-5914. DOI: 10.1111/j.1468-5914.1996.tb00285.x. URL: http://onlinelibrary.wiley.com.proxy.uchicago.edu/doi/10.1111/j.1468-5914.1996.tb00285.x/abstract (visited on 01/20/2012).

[124] Jens Høyrup. *Lengths, widths, surfaces: a portrait of old Babylonian algebra and its kin*. Springer, 2002 (cit. on p. 5).

[125] Frank Jackson. “Mental Causation”. In: *Mind* 105.419 (July 1996), pp. 377–413. URL: http://www.jstor.org/stable/2254828 (visited on 12/08/2013) (cit. on p. 12).

[126] Michael Kane. “All Validity Is Construct Validity. Or Is It?” In: *Measurement: Interdisciplinary Research & Perspective* 10.1-2 (2012), pp. 66–70. ISSN: 1536-6367. DOI: 10.1080/15366367.2012.681977.

[127] Antti Kauppinen. “The Rise and Fall of Experimental Philosophy”. In: *Philosophical Explorations* 10.2 (2007), pp. 95–118. DOI: 10.1080/13869790701305871.

[128] Matthias Kiesselbach. “Constructing Commitment: Brandom’s Pragmatist Take on Rule-Following”. en. In: *Philosophical Investigations* 35.2 (2012), pp. 101–126. DOI: 10.1111/j.1467-9205.2011.01450.x. (Visited on 12/14/2012).

[129] Gary King et al. “Enhancing the Validity and Cross-cultural Comparability of Measurement in Survey Research”. In: *American Political Science Review* 98 (2004), pp. 191–207. DOI: 10.1007/978-3-531-91826-6_16.

[130] Suzanne R. Kirschner. “Sociocultural Subjectivities Progress, Prospects, Problems”. In: *Theory & Psychology* 20.6 (Dec. 2010), pp. 765–780. ISSN: 0959-3543, 1461-7447. DOI: 10.1177/0959354310375745. URL: http://tap.sagepub.com.proxy.uchicago.edu/content/20/6/765 (visited on 01/23/2013).

[131] Paul Kline. “Commentary on Michell, Quantitative Science and the definition of measurement in psychology”. In: *British Journal of Psychology* 88.3 (1997), pp. 358–387. DOI: 10.1111/j.2044-8295.1997.tb02642.x.

[132] Joshua Knobe. “Experimental Philosophy and Philosophical Significance”. In: *Philosophical Explorations* 10.2 (2007), pp. 119–121. DOI: 10.1080/13869790701305905.

[133] Joshua Knobe. *Experiments in Philosophy*. http://opinionator.blogs.nytimes.com/2010/09/07/experimental-philosophy/. Sept. 7, 2010. URL: http://opinionator.blogs.nytimes.com/2010/09/07/experimental-philosophy/ (visited on 12/24/2012).

[134] Joshua Knobe et al. “Experimental Philosophy”. In: *Annual Review of Psychology* 63.1 (2012), pp. 81–99. DOI: 10.1146/annurev-psych-120710-100350.

[135] Robert Kraut. “Varieties of Pragmatism”. In: *Mind* 99.394 (Apr. 1990), pp. 157–183. URL: http://www.jstor.org/stable/2254958 (visited on 12/08/2013).

[136] Frederick Kroon. “Theory-dependence, warranted reference, and the epistemic dimensions of realism”. In: *Euro Jnl Phil Sci* 1.2 (May 2011), pp. 173–191. DOI: 10.1007/s13194-010-0004-4.

[137] Donald Laming. “A critique of a measurement–theoretic critique: Commentary on Michell, Quantitative science and the definition of measurement in psychology”. In: *British Journal of Psychology* 88.3 (1997), pp. 389–391. DOI: 10.1111/j.2044-8295.1997.tb02643.x.

[138] Suzanne Lane. “Consequences of Assessment and Accountability Systems Are Integral to the Argument-Based Approach to Validity”. In: *Measurement: Interdisciplinary Research & Perspective* 10.1-2 (2012), pp. 71–74. DOI: 10.1080/15366367.2012.679160. (Visited on 01/24/2013).

[139] Jeanette A Lawrence and Agnes E Dodds. “Three approaches to the disjunction between psychological measurement and psychological persons: methodological and ethical considerations”. In: *Integr Psychol Behav Sci* 44.4 (Dec. 2010). PMID: 20490955, pp. 299–309. DOI: 10.1007/s12124-010-9129-4.

[140] David Lewis. “Psychophysical and theoretical identifications”. In: *Australasian Journal of Philosophy* 50.3 (1972), pp. 249–258. DOI: 10.1080/00048407212341301.

[141] Robert W. Lissitz and Tiago Caliço. “Validity is an Action Verb: Commentary on: “Clarifying the Consensus Definition of Validity””. In: *Measurement: Interdisciplinary Research & Perspective* 10.1-2 (2012), pp. 75–79. DOI: 10.1080/15366367.2012.677346.

[142] Robert W. Lissitz and Karen Samuelsen. “A Suggested Change in Terminology and Emphasis regarding Validity and Education”. In: *Educational Researcher* 36.8 (Nov. 2007), pp. 437–448. DOI: 10.3102/0013189X07311286.

[143] Robert W. Lissitz and Karen Samuelsen. “Further Clarification regarding Validity and Education”. In: *Educational Researcher* 36.8 (Nov. 2007), pp. 482–484. DOI: 10.2307/4621104. URL: http://www.jstor.org/stable/4621104 (visited on 01/27/2013).

[144] Ronald Loeffler. “Neo-Pragmatist (Practice-Based) Theories of Meaning”. In: *Philosophy Compass* 4.1 (2009), pp. 197–218. DOI: 10.1111/j.1747-9991.2008.00186.x.

[145] Nathalie Lovasz and Kathleen L. Slaney. “What makes a hypothetical construct “hypothetical”? Tracing the origins and uses of the ‘hypothetical construct’ concept in psychological science”. In: *New Ideas in Psychology* 31.1 (Apr. 2013), pp. 22–31. DOI: 10.1016/j.newideapsych.2011.02.005.

[146] A. D. Lovie. “Commentary on Michell, Quantitative science and the definition of measurement in psychology”. In: *British Journal of Psychology* 88.3 (1997), pp. 393–394. DOI: 10.1111/j.2044-8295.1997.tb02644.x.

[147] E. J. Lowe. “Non-Cartesian Substance Dualism and the Problem of Mental Causation”. In: *Erkenntnis (1975-)* 65.1 (Jan. 2006), pp. 5–23. URL: http://www.jstor.org/stable/27667849 (visited on 12/08/2013) (cit. on p. 12).

[148] E. J. Lowe. “The Causal Autonomy of the Mental”. In: *Mind* 102.408 (Oct. 1993), pp. 629–644. URL: http://www.jstor.org/stable/2254049 (visited on 12/08/2013) (cit. on p. 12).

[149] R. Duncan Luce. “Quantification and symmetry: Commentary on Michell, Quantitative science and the definition of measurement in psychology”. In: *British Journal of Psychology* 88.3 (1997), pp. 395–398. DOI: 10.1111/j.2044-8295.1997.tb02645.x.

[150] R. Duncan Luce and Patrick Suppes. “Representational Measurement Theory”. In: *Stevens’ Handbook of Experimental Psychology*. John Wiley & Sons, Inc., 2002. URL: http://onlinelibrary.wiley.com.proxy.uchicago.edu/doi/10.1002/0471214426.pas0401/abstract (visited on 01/27/2013).

[151] William Lycan. “Representational Theories of Consciousness”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Fall 2008. 2008. URL: http://plato.stanford.edu/archives/fall2008/entries/consciousness-representational/.

[152] Michael Lynch. “Cognitive activities without cognition? ethnomethodological investigations of selected ‘cognitive’ topics”. In: *Discourse Studies* 8.1 (Feb. 2006), pp. 95–104. DOI: 10.1177/1461445606059559.

[153] Michael Lynch. “Ethnomethodology and the logic of practice”. In: *The Practice Turn in Contemporary Theory*. Routledge, 2001, pp. 140–157.

[154] Michael Lynch. “Method: Measurement - ordinary and scientific measurement as ethnomethodological phenomena”. In: *Ethnomethodology and the Human Sciences*. Cambridge Univ. Press, 1991, pp. 77–108.

[155] Cynthia MacDonald and Graham MacDonald. “Mental Causes and Explanation of Action”. In: *The Philosophical Quarterly* 36.143 (Apr. 1986). ArticleType: research-article / Issue Title: Special Issue: Mind, Causation and Action / Full publication date: Apr., 1986 / Copyright © 1986 The Philosophical Quarterly, pp. 145–158. DOI: 10.2307/2219765 (cit. on p. 12).

[156] Graham Macdonald and Philip Pettit. “The Nature of Naturalism”. In: *Proceedings of the Aristotelian Society, Supplementary Volumes* 66 (Jan. 1992), pp. 225–266. ISSN: 0309-7013. URL: http://www.jstor.org/stable/4106978 (visited on 12/08/2013).

[157] Alan R. Malachowski. *The New Pragmatism*. Acumen, 2010.

[158] Norman Malcolm. “Wittgenstein: The relation of language to instinctive behaviour”. In: *Philosophical Investigations* 5.1 (Jan. 1982), pp. 3–22. DOI: 10.1111/j.1467-9205.1982.tb00531.x.

[159] Michael D. Maraun. “Validity and Measurement”. In: *Measurement: Interdisciplinary Research & Perspective* 10.1-2 (2012), pp. 80–83. ISSN: 1536-6367. DOI: 10.1080/15366367.2012.682523.

[160] Michael D. Maraun and Stephanie M. Gabriel. “Illegitimate concept equating in the partial fusion of construct validation theory and latent variable modeling”. In: *New Ideas in Psychology* 31.1 (Apr. 2013), pp. 32–42. DOI: 10.1016/j.newideapsych.2011.02.006.

[161] Joseph Margolis. *Pragmatism without foundations: reconciling realism and relativism*. 2nd ed. London ; New York: Continuum, 2007. ISBN: 9780826491374.

[162] Joseph Margolis. *Reinventing pragmatism: American philosophyat the end of the 20th century*. Cornell University Press, 2002. ISBN: 9780801439957.

[163] S.W. Margulis. “Measurement Error and Reliability”. In: *Encyclopedia of Animal Behavior*. Ed. by Editors-in-Chief: Michael D. Breed and Janice Moore. Oxford: Academic Press, 2010, pp. 424–428. URL: http://www.sciencedirect.com/science/article/pii/B978008045337800231X.

[164] Keith A. Markus. “Constructs and Attributes in Test Validity: Reflections on Newton’s Account”. In: *Measurement: Interdisciplinary Research & Perspective* 10.1-2 (2012), pp. 84–87. DOI: 10.1080/15366367.2012.677348.

[165] Keith A. Markus. “Science, Measurement, and Validity: Is Completion of Samuel Messick’s Synthesis Possible?” In: *Social Indicators Research* 45.1/3 (Nov. 1998), pp. 5–34. DOI: 10.1023/A:1006960823277.

[166] Keith A. Markus and Denny Borsboom. “Reflective measurement models, behavior domains, and common causes”. In: *New Ideas in Psychology* 31.1 (Apr. 2013), pp. 54–64. DOI: 10.1016/j.newideapsych.2011.02.008.

[167] Keith A. Markus and Denny Borsboom. “The cat came back: Evaluating arguments against psychological measurement”. In: *Theory & Psychology* 22.4 (Aug. 2012), pp. 452–466. DOI: 10.1177/0959354310381155.

[168] Aryn Martin and Michael Lynch. “Counting Things and People: The Practices and Politics of Counting”. In: *Social Problems* 56.2 (May 2009), pp. 243–266. URL: http://www.jstor.org/stable/10.1525/sp.2009.56.2.243 (visited on 10/16/2012) (cit. on pp. 16, 29).

[169] Jack Martin. “Positivism, Quantification and the Phenomena of Psychology”. In: *Theory & Psychology* 13.1 (Feb. 2003), pp. 33–38. DOI: 10.1177/0959354303013001760.

[170] Per Martin-Löf. “Truth of a proposition, evidence of a judgement, validity of a proof”. In: *Synthese* 73.3 (Dec. 1987), pp. 407–420. DOI: 10.1007/BF00484985.

[171] Rosa L Matzkin. “Nonparametric Survey Response Errors”. In: *International Economic Review* 48.4 (Nov. 2007), pp. 1411–1427. DOI: 10.1111/j.1468-2354.2007.00468.x.

[172] Douglas W. Maynard and Steven E. Clayman. “The Diversity of Ethnomethodology”. In: *Annual Review of Sociology* 17.1 (1991), pp. 385–418. DOI: 10.1146/annurev.so.17.080191.002125. URL: http://www.annualreviews.org/doi/abs/10.1146/annurev.so.17.080191.002125 (visited on 02/18/2013).

[173] Douglas W. Maynard and Nora Cate Schaeffer. “Toward a Sociology of Social Scientific Knowledge: Survey Research and Ethnomethodology’s Asymmetric Alternates”. In: *Social Studies of Science* 30.3 (June 2000). ArticleType: research-article / Full publication date: Jun., 2000 / Copyright © 2000 Sage Publications, Ltd., pp. 323–370. ISSN: 0306-3127. DOI: 10.2307/285806. URL: http://www.jstor.org/stable/285806 (visited on 01/30/2013).

[174] Roderick P. McDonald. “Measuring Latent Quantities”. In: *Psychometrika* 76.4 (Oct. 2011), pp. 511–536. DOI: 10.1007/s11336-011-9223-7.

[175] Louis Menand. “An Introduction to Pragmatism”. In: *Pragmatism: A Reader*. Ed. by Louis Menand. New York: Vintage Books, 1997.

[176] Louis Menand, ed. *Pragmatism: a Reader*. New York: Vintage Books, 1997.

[177] Peter Menzies and Huw Price. “Causation as a Secondary Quality”. In: *The British Journal for the Philosophy of Science* 44.2 (June 1993). ArticleType: research-article / Full publication date: Jun., 1993 / Copyright © 1993 The British Society for the Philosophy of Science, pp. 187–203. ISSN: 0007-0882. URL: http://www.jstor.org/stable/687643 (visited on 12/08/2013) (cit. on p. 12).

[178] Samuel Messick. “Test Validity: A Matter of Consequence”. In: *Social Indicators Research* 45.1-3 (Nov. 1998), pp. 35–44. DOI: 10.1023/A:1006964925094.

[179] Samuel Messick. “Validity of psychological assessment: Validation of inferences from persons’ responses and performances as scientific inquiry into score meaning”. In: *American Psychologist* 50.9 (1995), pp. 741–749. ISSN: 1935-990X(Electronic);0003-066X(Print). DOI: 10.1037/0003-066X.50.9.741.

[180] Joel Michell. “Constructs, inferences, and mental measurement”. In: *New Ideas in Psychology* 31.1 (Apr. 2013), pp. 13–21. DOI: 10.1016/j.newideapsych.2011.02.004.

[181] Joel Michell. “Item Response Models, Pathological Science and the Shape of Error Reply to Borsboom and Mellenbergh”. In: *Theory & Psychology* 14.1 (Feb. 2004), pp. 121–129. DOI: 10.1177/0959354304040201.

[182] Joel Michell. “Normal Science, Pathological Science and Psychometrics”. In: *Theory & Psychology* 10.5 (Oct. 2000), pp. 639–667. DOI: 10.1177/0959354300105004 (cit. on pp. 16, 29).

[183] Joel Michell. “Quantitative science and the definition of measurement in psychology”. In: *British Journal of Psychology* 88.3 (1997), pp. 355–383. DOI: 10.1111/j.2044-8295.1997.tb02641.x (cit. on pp. 16, 29).

[184] Joel Michell. ““The constantly recurring argument”: Inferring quantity from order”. In: *Theory & Psychology* 22.3 (June 2012), pp. 255–271. DOI: 10.1177/0959354311434656.

[185] Joel Michell. “The psychometricians’ fallacy: Too clever by half?” In: *British Journal of Mathematical and Statistical Psychology* 62.1 (2009), pp. 41–55. DOI: 10.1348/000711007X243582.

[186] Joel Michell. “The Quantitative Imperative Positivism, Naive Realism and the Place of Qualitative Methods in Psychology”. In: *Theory & Psychology* 13.1 (Feb. 2003), pp. 5–31. DOI: 10.1177/0959354303013001758.

[187] Cheryl Misak, ed. *New Pragmatists*. Oxford University Press, USA, May 2007.

[188] Robert J. Mislevy. “The Case for Informal Argument”. In: *Measurement: Interdisciplinary Research & Perspective* 10.1-2 (2012), pp. 93–96. DOI: 10.1080/15366367.2012.682525. (Visited on 01/24/2013).

[189] Robert J. Mislevy. “Validity by Design”. In: *Educational Researcher* 36.8 (Nov. 2007), pp. 463–469. DOI: 10.3102/0013189X07311660.

[190] Hedwig te Molder and Jonathan Potter, eds. *Conversation and Cognition*. English. Cambridge; New York: Cambridge University Press, 2005.

[191] Michael Morgan. “Measurement in psychology: Commentary on Michell’s Quantitative Science and the definition of measurement in psychology”. In: *British Journal of Psychology* 88.3 (1997), pp. 399–400. DOI: 10.1111/j.2044-8295.1997.tb02646.x.

[192] Michael Morris. “Causes of Behaviour”. In: *The Philosophical Quarterly* 36.143 (Apr. 1986), pp. 123–144. DOI: 10.2307/2219764 (cit. on p. 12).

[193] Pamela A. Moss. “Reconstructing Validity”. In: *Educational Researcher* 36.8 (Nov. 2007), pp. 470–476. DOI: 10.3102/0013189X07311608.

[194] Pamela A. Moss. “Recovering a Dialectical View of Rationality”. In: *Social Indicators Research* 45.1/3 (Nov. 1998), pp. 55–67. DOI: 10.1023/A:1006925226003.

[195] Kevin R. Murphy. “Validity for What? The Peril of Overclarifying”. In: *Measurement: Interdisciplinary Research & Perspective* 10.1-2 (2012), pp. 97–99. DOI: 10.1080/15366367.2012.677362.

[196] Thomas Nadelhoffer and Eddy Nahmias. “The Past and Future of Experimental Philosophy”. In: *Philosophical Explorations* 10.2 (2007), pp. 123–149. ISSN: 1386-9795. DOI: 10.1080/13869790701305921.

[197] Paul E. Newton. “Clarifying the Consensus Definition of Validity”. In: *Measurement: Interdisciplinary Research & Perspective* 10.1-2 (2012), pp. 1–29. DOI: 10.1080/15366367.2012.669666.

[198] Paul E. Newton. “Questioning the Consensus Definition of Validity”. In: *Measurement: Interdisciplinary Research & Perspective* 10.1-2 (2012), pp. 110–122. DOI: 10.1080/15366367.2012.688456.

[199] Hugh J. Parry and Helen M. Crossley. “Validity of Responses to Survey Questions”. In: *The Public Opinion Quarterly* 14.1 (Apr. 1950), pp. 61–80. DOI: 10.1086/266150.

[200] Elazar J. Pedhazur and Liora Pedhazur Schmelkin. *Measurement, design, and analysis: an integrated approach*. Psychology Press, 1991. ISBN: 9780805810639.

[201] Charles Sanders Peirce. “How to Make Our Ideas Clear”. In: *Popular Science Monthly* 12 (1878), pp. 286–302.

[202] David Pitt. “Mental Representation”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2012. 2012. URL: http://plato.stanford.edu/archives/win2012/entries/mental-representation/.

[203] Alastair Pollitt. “Validity Cannot Be Created, It Can Only Be Lost”. In: *Measurement: Interdisciplinary Research & Perspective* 10.1-2 (2012), pp. 100–103. DOI: 10.1080/15366367.2012.686868. (Visited on 01/24/2013).

[204] Deborah Poole. “Routine Testing Practices and the Linguistic Construction of Knowledge”. In: *Cognition and Instruction* 12.2 (Jan. 1994), pp. 125–150. DOI: 10.1207/s1532690xci1202_3.

[205] Jonathan Potter. “Post-Cognitive Psychology”. en. In: *Theory & Psychology* 10.1 (Feb. 2000), pp. 31–37. DOI: 10.1177/0959354300010001596.

[206] Dag Prawitz. “Inference and Knowledge”. In: *Logica Yearbook 2008*. London: College Publications, 2009, pp. 183–200.

[207] Dag Prawitz. “The epistemic significance of valid inference”. In: *Synthese* (Mar. 2011), pp. 1–12. DOI: 10.1007/s11229-011-9907-7.

[208] Dag Prawitz. “Truth as an Epistemic Notion”. In: *Topoi* (), pp. 1–8. DOI: 10.1007/s11245-011-9107-6.

[209] Dag Prawitz. *Validity of Inference*. Abstract. Feb. 2008.

[210] Huw Price. "'Could a Question be True?': Assent and the Basis of Meaning". In: *The Philosophical Quarterly* 33.133 (Oct. 1983), p. 354. DOI: 10.2307/2219162.

[211] Huw Price. *Expressivism, pragmatism and representationalism*. Cambridge, UK: Cambridge University Press, 2013.

[212] Huw Price. "Naturalism without representationalism". In: *Expressivism, Pragmatism, and Representationalism*. Cambridge: Cambridge University Press, 2013, pp. 3–21.

[213] Huw Price. "Pluralism, 'world', and the primacy of science". In: *Expressivism, Pragmatism, and Representationalism*. Cambridge: Cambridge University Press, 2013, pp. 22–44.

[214] Huw Price. "Two expressivist programmes, two bifurcations". In: *Expressivism, Pragmatism, and Representationalism*. Cambridge: Cambridge University Press, 2013, pp. 22–44.

[215] George Psathas. *Everyday language: studies in ethnomethodology*. Irvington Publishers : distributed by Halsted Press, 1979. ISBN: 9780470266700.

[216] Hilary Putnam. *Representation and reality*. Cambridge, Mass.: MIT Press, 1991. ISBN: 0262660741 9780262660747.

[217] Hilary Putnam. *The Collapse of the Fact/Value Dichotomy and Other Essays*. Cambridge, MA: Harvard University Press, 2002. ISBN: 0674009053.

[218] Hilary Putnam. "The Three Enlightenments". In: *After Cognitivism: A Reassessment of Cognitive Science and Philosophy*. Springer, 2009.

[219] Louis Quéré and Cédric Terzi. "Some Features of Pragmatist Thought Still Remain Insufficiently Explored in Ethnomethodology". In: *Qualitative Sociology* 34.1 (Nov. 2010), pp. 271–275. DOI: 10.1007/s11133-010-9182-9.

[220] Panu Raatikainen. "Causation, Exclusion, and the Special Sciences". In: *Erkenn* 73.3 (Nov. 2010), pp. 349–363. DOI: 10.1007/s10670-010-9236-0.

[221] Anne Warfield Rawls. "Garfinkel, Ethnomethodology and the Defining Questions of Pragmatism". In: *Qualitative Sociology* 34.1 (Dec. 2010), pp. 277–282. DOI: 10.1007/s11133-010-9185-6.

[222] François Recanati, Gerhard Preyer, and Georg Peter. "Literalism and Contextualism: Some Varieties". In: *Contextualism in Philosophy: Knowledge, Meaning and Truth*. Clarendon Press, 2005, pp. 171–196.

[223] Mark D. Reckase. "The Interaction of Values and Validity Assessment: Does a Test's Level of Validity Depend on a Researcher's Values?" In: *Social Indicators Research* 45.1-3 (Nov. 1998), pp. 45–54. DOI: 10.1023/A:1006973109164.

[224] A. Kimball Romney, Susan C. Weller, and William H. Batchelder. "Culture as Consensus: A Theory of Culture and Informant Accuracy". In: *American Anthropologist*. New Series 88.2 (June 1986), pp. 313–338. DOI: 10.1525/aa.1986.88.2.02a00020.

[225] Richard Rorty. “Method, Social Science, and Social Hope”. In: *Canadian Journal of Philosophy* 11.4 (Dec. 1981), pp. 569–588. DOI: 10.1080/00455091.1981.10716323. URL: http://www.tandfonline.com/doi/abs/10.1080/00455091.1981.10716323.

[226] Richard Rorty. *Philosophy and the Mirror of Nature*. Princeton, N.J.: Princeton University Press, 2009.

[227] Richard Rorty. “Representation, Social Practise, and Truth”. In: *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 54.2 (Sept. 1988), pp. 215–228. DOI: 10.1007/BF00354514.

[228] Richard Rorty. “The Philosopher as Expert”. In: *Philosopy and the Mirror of Nature*. Thirtieth Anniversary Edition. Princeton University Press, 2009, pp. 395–421.

[229] Philip J. Rosenbaum and Jaan Valsiner. “The un-making of a method: From rating scales to the study of psychological processes”. In: *Theory & Psychology* 21.1 (Feb. 2011), pp. 47–65. DOI: 10.1177/0959354309352913.

[230] Nora Ruck and Thomas Slunecko. “A Stale Challenge to the Philosophy of Science”. In: *Integr. psych. behav.* 44.2 (June 2010), pp. 168–175. DOI: 10.1007/s12124-010-9121-z.

[231] Gilbert Ryle. *The concept of mind*. University of Chicago Press ed. Chicago: University of Chicago Pres, 1984. ISBN: 0226732959.

[232] Uwe Saint-Mont. “What measurement is all about”. In: *Theory & Psychology* 22.4 (Aug. 2012), pp. 467–485. DOI: 10.1177/0959354311429997.

[233] Robert E. Sanders. “Validating ‘observations’ in discourse studies: A methodological reason for attention to cognition”. English. In: *Conversation and Cognition*. Ed. by Hedwig te Molder and Jonathan Potter. Cambridge; New York: Cambridge Univ. Press, 2005, pp. 57–78.

[234] Kevin Scharp. “Brandom on Communication”. In: *Philosophical Profiles in the Theory of Communication*. McGill-Queen’s University Press.

[235] Theodore R. Schatzki, Karin Knorr Cetina, and Eike von Savigny, eds. *The Practice Turn in Contemporary Theory*. Routledge, Jan. 2001.

[236] Annette C Scherpenzeel and Willem E Saris. “The Validity and Reliability of Survey Questions A Meta-Analysis of MTMM Studies”. In: *Sociological Methods & Research* 25.3 (Feb. 1997), pp. 341–383. DOI: 10.1177/0049124197025003004.

[237] Verena D. Schmittmann et al. “Deconstructing the construct: A network perspective on psychological phenomena”. In: *New Ideas in Psychology* 31.1 (Apr. 2013), pp. 43–53. DOI: 10.1016/j.newideapsych.2011.02.007.

[238] Peter Schroeder-Heister. “Validity Concepts in Proof-theoretic Semantics”. In: *Synthese* 148.3 (Feb. 2006), pp. 525–571. DOI: 10.1007/s11229-004-6296-1.

[239] Michael Schwarz. “Is Psychology Based on a Methodological Error?” In: *Integrative Psychological and Behavioral Science* 43.3 (Sept. 2009), pp. 185–213. DOI: 10.1007/s12124-009-9089-8 (cit. on pp. 17, 30).

[240] Wilfrid Sellars. *Empiricism and the Philosophy of Mind*. Cambridge, Mass.: Harvard University Press, 1997.

[241] Wilfrid Sellars. “Language as Thought and as Communication”. In: *Philosophy and Phenomenological Research* 29.4 (June 1969), pp. 506–527. DOI: 10.2307/2105537.

[242] David Sherry. “Thermoscopes, thermometers, and the foundations of measurement”. In: *Studies In History and Philosophy of Science Part A* 42.4 (Dec. 2011), pp. 509–524. DOI: 10.1016/j.shpsa.2011.07.001 (cit. on pp. 16, 29).

[243] Klaas Sijtsma. “Correcting Fallacies in Validity, Reliability, and Classification”. In: *International Journal of Testing* 9.3 (2009), pp. 167–194. DOI: 10.1080/15305050903106883.

[244] Klaas Sijtsma. “Future of Psychometrics: Ask What Psychometrics Can Do for Psychology”. In: *Psychometrika* 77.1 (Jan. 2012), pp. 4–20. DOI: 10.1007/s11336-011-9242-4.

[245] Klaas Sijtsma. “Psychological measurement between physics and statistics”. In: *Theory & Psychology* 22.6 (Dec. 2012), pp. 786–809. DOI: 10.1177/0959354312454353 (cit. on pp. 15, 28).

[246] Stephen G. Sireci. “On Validity Theory and Test Validation”. In: *Educational Researcher* 36.8 (Nov. 2007), pp. 477–481. DOI: 10.3102/0013189X07311609.

[247] Stephen G. Sireci. “The Construct of Content Validity”. In: *Social Indicators Research* 45.1-3 (Nov. 1998), pp. 83–117. DOI: 10.1023/A:1006985528729.

[248] Tom W. Smith. “Refining the Total Survey Error Perspective”. In: *International Journal of Public Opinion Research* 23.4 (2011), pp. 464–484. DOI: 10.1093/ijpor/edq052 (cit. on pp. 17, 30).

[249] Ernest Sosa. “Intuitions and meaning divergence”. In: *Philosophical Psychology* 23.4 (2010), pp. 419–426. DOI: 10.1080/09515089.2010.505859.

[250] Henderikus J. Stam. “The Fault is Not in Ourselves, but in Our Methods: Comment on Schwarz”. In: *Integrative Psychological and Behavioral Science* 44.4 (Dec. 2010), pp. 281–287. DOI: 10.1007/s12124-010-9144-5 (cit. on pp. 17, 30).

[251] S. S. Stevens. “On the Theory of Scales of Measurement”. In: *Science* 103.2684 (June 1946). PMID: 17750512, pp. 677–680. DOI: 10.1126/science.103.2684.677 (cit. on pp. 14, 27).

[252] R. Steyer. “Classical (Psychometric) Test Theory”. In: *International Encyclopedia of the Social & Behavioral Sciences*. Ed. by Editors-in-Chief: Neil J. Smelser and Paul B. Baltes. Oxford: Pergamon, 2001, pp. 1955–1962. DOI: 10.1016/B0-08-043076-7/00721-X. (Visited on 02/18/2013).

[253] Milton E. Strauss and Gregory T. Smith. “Construct Validity: Advances in Theory and Methodology”. In: *Annual Review of Clinical Psychology* 5.1 (2009). PMID: 19086835, pp. 1–25. DOI: 10.1146/annurev.clinpsy.032408.153639.

[254] B. G Sundholm. “Inference versus Consequence”. In: *LOGICA Yearbook 1997, 26 - 35 (1998)*. Prague: Filosofia Publishers, Czech Academy of Science, 1998, pp. 26–35. URL: https://openaccess.leidenuniv.nl/handle/1887/10419 (visited on 01/16/2012).

[255] Shirley Anne Tate. “Foucault, Bakhtin, Ethnomethodology: Accounting for Hybridity in Talk-in-Interaction”. In: *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* 8.2 (May 2007). ISSN: 1438-5627. URL: http://www.qualitative-research.net/index.php/fqs/article/view/247/545.

[256] L. L. Thurstone. “A Law of Comparative Judgment”. In: *Psychological Review* 34.4 (1927), pp. 273–286. DOI: 10.1037/h0070288.

[257] Michael Tomasello. *Origins of human communication*. 1st. MIT Press pbk. ed. Cambridge Mass. ;;London: MIT Press, 2010. ISBN: 9780262515207.

[258] Aaro Toomela. “Variables in Psychology: A Critique of Quantitative Psychology”. In: *Integrative Psychological and Behavioral Science* 42.3 (June 2008), pp. 245–265. DOI: 10.1007/s12124-008-9059-6 (cit. on pp. 17, 30).

[259] Roger Tourangeau, Lance J. Rips, and Kenneth A. Rasinski. *The Psychology of Survey Response*. Cambridge University Press, Mar. 2000. ISBN: 9780521576291.

[260] Günter Trendler. “Measurement Theory, Psychology and the Revolution That Cannot Happen”. In: *Theory & Psychology* 19.5 (Oct. 2009), pp. 579–599. DOI: 10.1177/0959354309341926.

[261] Anand J. Vaidya. “Philosophical methodology: The current debate”. In: *Philosophical Psychology* 23.3 (2010), pp. 391–417. ISSN: 0951-5089. DOI: 10.1080/09515089.2010.490940.

[262] Guy C. Van Orden, Christopher T. Kello, and John G. Holden. “Situated Behavior and the Place of Measurement in Psychological Theory”. In: *Ecological Psychology* 22.1 (2010), pp. 24–43. DOI: 10.1080/10407410903493145.

[263] Stéphane Vautier et al. “The ambiguous utility of psychometrics for the interpretative foundation of socially relevant avatars”. In: *Theory & Psychology* 22.6 (Dec. 2012), pp. 810–822. DOI: 10.1177/0959354312450093.

[264] Meike Watzlawik. ““Der Weisheits letzter Schluss”? Wisdom’s Last Conclusion?” In: *Integr. psych. behav.* 43.3 (Sept. 2009), pp. 214–220. DOI: 10.1007/s12124-009-9094-y.

[265] Bernhard Weiss and Jeremy Wanderer. *Reading Brandom: on making it explicit*. Taylor & Francis, June 2009. ISBN: 9780415380362.

[266] Michael A. Westerman. “Conversation analysis and interpretive quantitative research on psychotherapy process and problematic interpersonal behavior”. In: *Theory & Psychology* 21.2 (Apr. 2011), pp. 155–178. DOI: 10.1177/0959354310394719.

[267] Michael A. Westerman and Stephen C. Yanchar. “Changing the terms of the debate: Quantitative methods in explicitly interpretive research”. In: *Theory & Psychology* 21.2 (Apr. 2011), pp. 139–154. DOI: 10.1177/0959354310393565.

[268] Susan E. Whitely. “Construct validity: Construct representation versus nomothetic span”. In: *Psychological Bulletin* 93.1 (1983), pp. 179–197. DOI: 10.1037/0033-2909.93.1.179.

[269] Timothy Williamson. “The Broadness of the Mental: Some Logical Considerations”. In: *Noûs* 32 (Jan. 1998), pp. 389–410. URL: http://www.jstor.org/stable/2676155 (visited on 12/08/2013) (cit. on p. 12).

[270] Peter Winch. *The Idea of a Social Science and its Relation to Philosophy*. English. 2nd. London; New York: Routledge, 1958.

[271] Christopher Winship and Christopher Muller. “Ethnomethodology and Consequences: Comment on Emirbayer and Maynard’s “Pragmatism and Ethnomethodology””. In: *Qualitative Sociology* 34.1 (Nov. 2010), pp. 283–286. DOI: 10.1007/s11133-010-9179-4.

[272] Alan Wolfe. “The Missing Pragmatic Revival in American Social Science”. In: *The Revival of Pragmatism: New Essays on Social Thought, Law, and Culture*. Duke University Press, 1998.

[273] R. Yusoff and R. Mohd Janor. “Generation of an Interval Metric Scale to Measure Attitude”. In: *SAGE Open* 4.1 (Jan. 2014). ISSN: 2158-2440. DOI: 10.1177/2158244013516768. URL: http://sgo.sagepub.com/content/4/1/2158244013516768.full-text.pdf+html (visited on 01/16/2014).

[274] Annemarie Zand Scholten. “The Consensus Definition Redefined From a Representational Perspective”. In: *Measurement: Interdisciplinary Research & Perspective* 10.1-2 (2012), pp. 104–109. DOI: 10.1080/15366367.2012.681978.

[275] Don H. Zimmerman. “Review: Ethnomethodology and the Human Sciences”. In: *Journal of Linguistic Anthropology* 4.1 (1994), pp. 94–95. DOI: 10.1525/jlin.1994.4.1.94.

[276] Donald W. Zimmerman. “Commentary on ‘Science, Measurement, and Validity: Is Completion of Samuel Messick’s Synthesis Possible?’ by Keith A. Markus”. In: *Social Indicators Research* 45.1-3 (Nov. 1998), pp. 69–72. DOI: 10.1023/A:1006977210073.

[277] Donald W. Zimmerman. “How Should Classical Test Theory Have Defined Validity?” In: *Social Indicators Research* 45.1-3 (Nov. 1998), pp. 233–251. DOI: 10.1023/A:1006949915525.

[278] Bruno D. Zumbo. “Opening Remarks to the Special Issue on Validity Theory and the Methods Used in Validation: Perspectives from the Social and Behavioral Sciences”. In: *Social Indicators Research* 45.1/3 (Nov. 1998), pp. 1–3. DOI: 10.1023/A:1006956722368.