

V

validity, qualitative

Patti Lather

Validity is not just one of many issues in science but the crux of the issue: the claims of science to a certain privilege in terms of authoritative knowledge. How scientific knowledge is made credible is, hence, a longstanding issue. If one looks at validity as a social construction, one sees how the very calculus of credibility of what is deemed “good science,” the very determination of warrants of validity, has shifted across time, place, and various fields.

In the contemporary moment, the crisis of legitimation occurring across knowledge systems is registered in a cacophony of postpositivism, non-foundationalism, kinds of realism and post-realism, warranted assertability, logics of inquiry, construct validity, carefully controlled inference, objectivism, situational validity, and Cronbachian insights regarding the decay of generalizations. As a result, discourse practices of validity in qualitative research exemplify a proliferation of available framings in terms of the legitimation of knowledge, particularly the power and political dimensions of the issue of demarcation.

Various turns have characterized research in the human sciences over the last few decades, shifts that are not so much linear as multiple, simultaneous, and interruptive. It is as if the critiques of truth in Nietzsche, self-presence in Freud, referential language in Saussure, and metaphysics in Heidegger were finally coming home to roost in the social sciences. Across this dizzying array of in-movement shifts, these turns challenge the “view from nowhere” and the traditional foundations of knowledge that continue to undergird so much of contemporary research. The following outlines twentieth-century turns

toward epistemological indeterminacy in order to underscore contemporary interest in situatedness, perspective, relationality, narrative, poesis, and blurred genres. It then surveys across the field of social inquiry in terms of the variety of available discourses of validity in order to delineate the weakening of any “one best way approach” to validity.

EPISTEMIC INDETERMINACY AND THE WEAKENING OF HOMOGENEOUS STANDARDS

In exploring the work of science in an era of blurred genres, validity is a “limit question” of research, one that repeatedly resurfaces, one that can be neither avoided nor resolved. Within a context of epistemic anti-foundationalism, validity is about much more than the limits of objectivity: “It bores into the essence of science itself” (Kerlinger 1986: 432). What follows argues in a Foucauldian manner that validity be situated as practices toward spaces of constructed visibility and incitements to see which constitute power/knowledge. This post-epistemic focus decenters validity as about epistemological guarantees and shifts it into practices that are situated, multiple, partial, endlessly deferred, a reflexive validity interested in how discourse does its work.

From a post-epistemic focus, validity is a boundary line for what is acceptable and not acceptable in research. Validity is, in short, power, the power to determine the demarcation between science and not-science. Such a post-epistemic focus shifts the validity question in some interesting directions. Some argue for dismissing validity altogether as too much about the continuation of positivist ideals. Others worry that qualitative work that fails to provide systematic depth analysis and analytic rigor

threatens the fragile legitimacy that qualitative research has established and holds it to a scientific accounting. This is made most obvious in recent moves by the federal government to warrant experimental design as the “gold standard” for good science.

In contrast, Pam Moss, writing out of psychometrics and assessment, argues that all social science research is under theoretic pressure in terms of foundational assumptions. Moss (1996) argues for a reflexive complementarity between varied approaches to the social sciences in order to think reflexivity about our taken-for-granted practices and perspectives. Moss sees a reciprocity of accountability in this purposeful engagement across paradigmatic assumptions and her expansion of validity echoes Mishler’s (1990) argument that the “problem” of validity is about deep theoretical issues that technical solutions cannot begin to address. Ever since Cronbach and Meehl’s (1955) essay on the problems with construct validity in psychological testing, validity has been the problem, not the solution. Various post-positivist efforts have been made to resolve the problem, from the naturalistic and constructivist paradigms of Lincoln and Guba that dominated the early discourse of qualitative research to discourse theory, ethnographic authority, critical, feminist, and race-based paradigms and more recent poststructuralisms. Some efforts toward validity in qualitative research remain deeply inscribed in a correspondence model of truth and assumptions of transparent narration, while others attempt validity practices that take into account the crisis of representation. And some call for new imaginaries altogether, where validity is as much about the play of difference as the repetition of sameness. Rather than exhausting the problem, all exemplify how the effort to answer the problem of validity is always partial, situated, temporary.

The following traces these provisional “solutions” as an effort to displace normative criteria of quality. Normative criteria posit themselves as universal and attempt to regulate “best way” procedures, whereas socially grounded criteria are situated, relational, temporal/historical. Unlike standardized regulatory criteria, such criteria move away from compelling conviction to some essence and toward contextually relevant practices that both disrupt referential logic and shift orientation from the object to the

relations of its perception, to its situation of address and reception.

COUNTERPRACTICES OF AUTHORITY: FROM QUALITY CRITERIA TO SOCIAL PRACTICES

Just a decade ago, Lincoln and Guba’s delineation of validity served as a sort of mantra across qualitative work in the social sciences. This evidences the importance of a validity discourse appropriate to qualitative research, but most interesting is how Guba and Lincoln’s early delineation worked in unanticipated ways to undercut representational logic and spawn increasingly post-epistemic practices of validity. This section traces the movement of their thinking across a decade of validity formulations. To set the stage for this, the first layer in the story of validity in qualitative research is the standard story from the side of positivism.

Whereas the criteria for the credibility of quantitative research are based on the validity and reliability of instruments and internal validity, in qualitative research the primary criterion is the credibility of the study. Credibility is defined as the extent to which the data, data analysis, and conclusions are believable and trustworthy as based on a set of standard practices. Markers of credibility include triangulation, the use of different methods, samples of people and/or times and places. Reliability is the fit between what occurs and what is recorded, and is established by: detailed fieldnotes, a team approach, participant confirmation of accuracy of observations, mechanized recording of data (tape recorders, videotapes, photographs), use of participant quotations, and an active search for discrepant data. Internal validity refers to the match between researchers’ categories and interpretations and what is actually true. It is claimed via prolonged engagement, thick description, thorough delineation of research process, and unobtrusive entry and participation in the setting. Finally, external validity shifts from generalizability based on sampling to reader assessment of transferability.

While this treatment of generalization evidences some attention to post-positivist assumptions, the preceding is grounded in the sort of scientificity that is at issue here. Guba and

Lincoln (1989), for example, argue that internal validity, as an assessment of the degree of isomorphism between a study's findings and the real world, cannot have meaning as a criterion in a paradigm that rejects a realist ontology. Additionally, external validity or generalizability has little meaning if realities are multiple and constructed. Erickson's (1986) idea of "particularizability" seems more useful: documenting particular cases with "thick" description, so that the reader can determine the degree of "transferability." Most interesting in this standard treatment of validity in qualitative research is the rather unremarked work of the concept of "transferability." Displacing a validity of correspondence with a focus on the terms of address, of reception, shifts orientation to the reader who determines the degree to which a study is "transferable" to his or her own context of interest.

The next layer in the story of validity is a standard treatment of validity from the side of post-positivism.

Michael Patton's *Qualitative Evaluation and Research Methods* (1990 [1980]) was one of the most widely used texts prior to the bestselling *Handbook of Qualitative Research* (Denzin & Lincoln 2000 [1994]). Patton elaborates on methodical reporting of systematic procedures of data collection and analysis. Particularly concerned about researcher effects, he cautions against the sort of self-importance that often leads to overrating this problem. The key is that reducing distortions is based on "empathic neutrality," a kind of impartiality that works to minimize researcher effect while recognizing that "the data inevitably represent perspective rather than absolute truth" (p. 475). In delineating legitimating practices, Patton surveys across the most frequently noted figures: Lincoln and Guba and Miles and Huberman on specific validity practices; LeCompte and Goetz and Kirk and Miller on reliability and validity; Michael Scriven on rethinking objectivity; Denzin on triangulation; Peshkin on subjectivity as a resource; and Cronbach on generalizability. The basic assumptions of this canonical discourse on validity in qualitative research can be traced by unpacking the work of, arguably, the central figures in the validity debates in qualitative research, Guba and Lincoln.

In the summary chart in *Naturalistic Inquiry* (1985), Lincoln and Guba summarize the techniques for establishing trustworthiness as (1) credibility (prolonged engagement and persistent observation; triangulation of sources, methods, and investigators; peer debriefing; negative case analysis; referential adequacy; and member checks); (2) transferability via thick description; (3) dependability and confirmability via an audit trail; and (4) the reflexive journal (p. 328). Each practice is more or less developed, with the member check positioned as the most crucial technique. This involves taking back to the participants what you have learned from them and can range from a minimalist "transcript check" to a more involved reaction to a preliminary analysis to a maximal feedback loop in regards to the final write-up. All are offered in the hopes of working against prescription and orthodoxy.

By 1989, Guba and Lincoln had moved to a delineation of three different approaches: parallel or quasi-foundational criteria, now called trustworthy criteria; the nature of the hermeneutic process itself; and a new set of non-foundational criteria, termed the authenticity criteria. The parallel criteria map onto the 1985 formulation, but they are more clearly located in a post-realist ontology, for example, triangulation is deemphasized as "too positivist" in its assumptions of "unchanging phenomena" (p. 240). "The hermeneutic process as its own quality control" argues the difficulty of falsity because of the interactive, dialogic nature of the research process. The most noteworthy feature of the authenticity criteria is the break with more traditional methodological criteria into criteria that blur the line between ethics and validity. Termed *fairness* and *ontological, educative, catalytic, and tactical authenticity*, the criteria are about balancing viewpoints, encouraging the learning of both researcher and researched, sharing knowledge democratically, and fostering social action. The emphasis here is the move of validity from a set of epistemic concepts to a space of relational practices in situated contexts of inquiry.

By 1995, Lincoln shifted fully into an anti-foundational discourse interested in research as relational and fostering of action and social justice. Quality criteria are posited as fluid and emergent, with a focus on criteria that collapse the distinction between rigor and ethics.

Tracing both the history and the rationale for the continued importance of rigor criteria, Lincoln notes her continued use of the parallel foundationalist criteria with her doctoral students as a place to begin. She then delineates emerging criteria that, while all relational, are differently aware of the exclusionary function of quality criteria and the inevitability of partial and incomplete standpoints. Regarding the latter, “detachment and author objectivity” become “barriers to quality, not insurance of having achieved it,” as she urges researchers to “come clean” about their own stances (p. 280). Epistemology is situated as an ethical issue, and objectivism is displaced by linking research as a community project to social action. Key practices are delineated: the use of multiple voices, reflexivity regarding the relationships and contradictions of research processes, reciprocity, sacredness, and sharing royalties as a way to address the cultural and economic capital that academics make out of the lives of others. This includes movement toward action inquiry. The interest here is the move beyond the search for uniform criteria toward criteria that emerge as a natural consequence of the inquiry effort. This is a call for a profusion of situated validities, immanent validities, within the context of a particular inquiry.

Seeing validity as an apparatus of betterment, as a cure for what ails us, Lincoln’s panegyric contrasts starkly with Scheurich (1996) who, rather than pay tribute, deconstructs “the masks of validity.” Across both positivism and post-positivism, Scheurich organizes discourses of validity into three categories. The first, *originary validity*, translates conventional science concerns into post-positivism, for example Lincoln and Guba’s parallel criteria. *Successor validity* recasts the concepts that arose in opposition to conventional notions of science, for example the concept of catalytic validity that grows out of advocacy research or “research as praxis” (Lather 1986a, b). Finally, *interrogated validity* deconstructs the policing function of validity, for example Cherryholmes’s argument that construct validity is “of and about power” (1988: 450). Scheurich argues that to the extent discourse practices of validity are about policing the borders between “the accepted from the not true or the unaccepted or the not yet accepted” (1996: 5), they are

“imperial” in allowing the same and disallowing the different. At the heart of the western knowledge project, Scheurich writes, is this “Same/Other power binary” (p. 6) that is more about “Eating the Other” (quoting bell hooks) than it is about increasing knowledge. “Validity practices are unconscious instantiations of a western philosophical . . . dualism” (p. 8) that is not about individual conscious intentions but about the western “civilizational project, an imperial project” (p. 7). To undermine this dualism, he urges new imaginaries of validity that both unmask dualisms and celebrate polyphony and difference, the shifting complexities of truth as multiply perspectival.

As a possibility, Scheurich unpacks Lather’s (1993) delineation of transgressive validities – ironic, paralogical, rhizomatic, and situated/embodied/voluptuous. All unsettle truth regimes, implode controlling codes, and work against the constraints of authority. All foreground the insufficiencies of language and the production of meaning effects, foster differences and heterogeneity, put conventional discursive procedures under erasure, and embody a situated, partial, positioned, explicit tentativeness. All anticipate a politics that desires both justice and the unknown, generate new locally determined norms of understanding, and proliferate open-ended and context-sensitive criteria that enact practices of engagement and self-reflexivity. All bring ethics and epistemology together. Intended to “incite” the proliferation of validity discourse practices, this effort leaves Scheurich unsatisfied, however, still worried about the capacity of “our restless civilizational immodesty” to reappear with new masks in its continuing absorption of the other into the same (1996: 10). Turning to the accelerating proliferation of marginalized voices, he calls for “a Bakhtinian dialogic carnival, a loud clamor of a polyphonic, open, tumultuous, subversive conversation on validity” (p. 10). Here validity has moved from a discourse about *quality* as normative to a discourse of *relational practices* that evokes an epistemic disruption, a transgression of set forms.

This exemplifies how validity is being used to further change the terms of the legitimation of knowledge beyond discrete methods and toward the social uses of the knowledge we

construct. Across shifts in episteme and the consequent weakening of homogeneous standards and the proliferation of counterpractices of authority in qualitative research, the intelligibility and availability of alternative discourse practices of validity work to loosen positivism and suggest the critical potential of validity to put under theoretic pressure the claims of scientificity.

SEE ALSO: Culture; Legitimacy; Methods; Reliability; Representation; Science, Social Construction of; Scientific Knowledge, Sociology of; Theory; Theory and Methods; Trustworthiness

REFERENCES AND SUGGESTED READINGS

- Cherryholmes, C. (1988) *Power and Criticism: Post-structural Investigations in Education*. Teachers College Press, New York.
- Cronbach, L. & Meehl, P. (1955) Construct Validity in Psychological Tests. *Psychological Bulletin* 52: 281–302.
- Denzin, N. & Lincoln, Y. (Eds.) (2000 [1994]) *Handbook of Qualitative Research*, 2nd edn. Sage, Thousand Oaks, CA.
- Erickson, F. (1986) Qualitative Methods in Research on Teaching. In: Wittrock, M. C. (Ed.), *Handbook of Research on Teaching*. Macmillan, New York, pp. 119–61.
- Guba, E. & Lincoln, Y. (1989) *Fourth-Generation Evaluation*. Sage, Newbury Park, CA.
- Kerlinger, F. (1986) *Foundations of Behavioral Research*, 3rd edn. Holt, Rinehart, & Winston, New York.
- Kvale, S. (1995) The Social Construction of Validity. *Qualitative Inquiry* 1(1): 19–40.
- Lather, P. (1986a) Research as Praxis. *Harvard Educational Review* 56(3): 257–77.
- Lather, P. (1986b) Issues of Validity in Openly Ideological Research: Between a Rock and a Soft Place. *Interchange* 17(4): 63–84.
- Lather, P. (1993) Fertile Obsession: Validity After Poststructuralism. *Sociological Quarterly* 34(4): 673–93.
- Lincoln, Y. (1995) Emerging Criteria for Quality in Qualitative and Interpretive Research. *Qualitative Inquiry* 1(3): 275–89.
- Lincoln, Y. & Guba, E. (1985) *Naturalistic Inquiry*. Sage, Newbury Park, CA.
- Mishler, E. (1990) Validation in Inquiry-Guided Research: The Role of Exemplars in Narrative Studies. *Harvard Educational Review* 60(4): 415–42.
- Moss, P. (1996) Enlarging the Dialogue in Educational Measurement: Voices from Interpretive Research Traditions. *Educational Researcher* (January/February): 20–8, 43.
- Patton, M. (1990 [1980]) *Qualitative Evaluation and Research Methods*. Sage, Newbury Park, CA.
- Scheurich, J. (1996) The Masks of Validity: A Deconstructive Investigation. *Qualitative Studies in Education* 9(1): 49–60.

validity, quantitative

Karen Lahm

The term validity can be defined and explained in a plethora of ways. In quantitative research, validity is most commonly discussed when a researcher is developing measures of variables or concepts to put on a survey or use in an experiment, etc. Specifically, validity is synonymous with accuracy in that a valid measure is one that is actually or in reality measuring the concept or variable that it is supposed to be measuring. For example, say a researcher needed to develop a measure for the variable IQ (Intelligence Quotient). So the researcher decides that he is going to ask his sample to get on a scale and weigh themselves in pounds. Such a measure of IQ (body weight in pounds) would be an invalid measure of one's intelligence because in reality body weight tells us nothing about one's IQ. A more valid measure of one's intelligence would be to give an IQ test and compute IQs from the answers gathered on the test.

In order to ensure validity, researchers strive to meet several criteria when they create their measures of concepts and/or variables. First, they try to ensure that their measure has face validity. Face validity provides a check to see if “on its face” or literally it is a measure appropriate for a concept. The example above lacks face validity because in reality body weight does not tell us anything about how intelligent one is. A second criterion for assessing measures is content validity. This suggests that a measure must cover a range of meanings for a given concept or variable. For example, if a researcher wanted to measure someone's criminality, he would not just ask them about how many times