

International Journal of Social Research Methodology

Visual Models, Piloting, and Triangulation as Formative Research Practices

Submission ID	255889033
Article Type	Research Article
Keywords	Formative Evaluation, Visual Model, Instrument Pilot, Analysis Pilot, Triangulation
Authors	Patrick Biddix

For any queries please contact:

TSRM-peerreview@journals.tandf.co.uk

Note for Reviewers:

To submit your review please visit <https://mc.manuscriptcentral.com/tsrm>

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18 **Visual Models, Piloting, and Triangulation as Formative Research Practices**
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38

39 **J. Patrick Biddix, Ph.D.**

40 Jimmy and Ileen Cheek Endowed Professor of Higher Education
41 Department of Educational Leadership and Policy Studies (ELPS)
42 College of Education, Health and Human Services (CEHHS)
43 University of Tennessee, Knoxville, TN, USA
44 1122 Volunteer Blvd., 316 Bailey Education Complex, Knoxville, TN 37996-3430
45 E-mail: pbiddix@utk.edu; Tel.: 1-865-974-2216

46 **ORCID ID**

47 J. Patrick Biddix <https://orcid.org/0000-0002-6977-9548>
48
49
50
51
52
53
54
55
56
57
58
59
60

Visual Models, Piloting, and Triangulation as Formative Research Practices

Researchers frequently encounter methodological challenges, or “messy areas” (Cook, 2009), requiring critical but often undocumented decisions about design, data collection, and analysis (Sandelowski, 1993; Gadella Kamstra, 2022). While validity and analogous concepts and strategies promote accuracy and transparency (Lincoln & Guba, 1985; Shenton, 2004), they can also feel restrictive (Van Assche et al., 2021) and conceptually problematic (Ong, 2024). To navigate these complexities, common responsive techniques are needed to monitor and adapt inquiry approaches and document the process. This article proposes four cross-paradigmatic practices—visual models, instrument pilots, analysis pilots, and triangulation—grounded in a formative evaluation approach (Black & Wiliam, 2009) to enhance rigor and quality in social science research.

Searching for Common Evaluative Research Practices

Evaluative practices in social science research vary by paradigm, epistemology, and methodology. Despite efforts to identify standard concepts, consensus remains elusive (Morse, 2015). Researchers have proposed equivalent terminology (Lincoln & Guba, 1985; Shenton, 2004), broad approaches (Barker & Pistrang, 2005), and conceptual frameworks (Dellinger & Leech, 2007), yet common practices to address invalidity or trustworthiness remain scarce (Johnson & Onwuegbuzie, 2004). Most recommendations rely on validity language (Maxwell, 2013; Ong, 2024) or method-specific practices (Shadish et al., 2002). Barker and Pistrang (2005) observed that, despite its contentious nature, pursuing common evaluation criteria could lead to “a new integration of existing ideas” (p. 203). Tashakkori et al. (2020) emphasized that

1
2
3
4
5
6
7
8 establishing clear concepts strengthens inferences by enabling researchers to defend
9 their conclusions and rule out alternative explanations. However, as Onwuegbuzie and
10 Leech (2005) noted, diverse research methodologies “are merely tools that are
11 designed to aid our understanding of the world” (p. 377), suggesting validity criteria
12 should be adaptable. Building on these discussions, validity serves as a conceptual
13 foundation, with rigor and quality as standards to identify common evaluative practices.
14
15
16
17
18
19

20 **Identifying Validity as a Conceptual Foundation**

21
22 Validity traditionally is defined as the degree to which a concept is accurately
23 measured or represented (Shadish et al., 2002). It is often seen as a contingent,
24 context-dependent construct rather than a fixed, universal standard (Maxwell, 2013). In
25 quantitative research, validity is linked to confidence, achieved by minimizing external
26 influences to isolate effects on outcomes (Fowler, 2014). Lincoln and Guba (1985)
27 criticized validity practices as deficit-based, arguing that by prioritizing control over
28 context, they limit relevance and distance researchers from data. They proposed
29 trustworthiness as analogous to validity, encompassing credibility (internal validity),
30 transferability (external validity), dependability (reliability), and confirmability (objectivity)
31 (see also Shenton, 2004). Sandelowski (1993) viewed validity and trustworthiness as
32 interconnected aspects of theory, emphasizing sound science over claiming accuracy.
33 This perspective aligns with Dellinger and Leech’s (2007) mixed methods evaluative
34 framework, which emphasized legitimation. As a conceptual foundation for this article,
35 validity is defined as ensuring accuracy to enhance relevance, meaning, and
36 appropriateness.
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Applying Formative Evaluation to Enhance Rigor and Quality

Gadella Kamstra (2022), echoing Cook (2009), emphasized that “presenting frank descriptions of the research process, embracing mess, and honestly reporting it as part of the investigation will ensure the research’s strength or validity” (p. 742). A **formative evaluation** approach is designed to monitor progress and identify areas for improvement (Black & Wiliam, 2009). Quantitative researchers often incorporate formative practices like pilot testing and data validation to refine methods and ensure reliability (Fowler, 2014). Qualitative researchers utilize iterative practices throughout the inquiry process (Patton, 2014), as they “move back and forth between design and implementation to ensure congruence among question formulation, literature, recruitment, data collection strategies, and analysis” (Morse et al, 2002, p. 17). Mixed methods researchers use conceptual frameworks and apply method-specific validity strategies during data collection and integration (Dellinger & Leech, 2007; Tashakkori et al., 2020).

A formative approach to **rigor** emphasizes accuracy, appropriateness, soundness, and consistency during the conceptual, planning, and testing phases of a study. Key actions for establishing rigor include setting clear goals, aligning goals with research questions, identifying relevant data, and applying sound analytic procedures. From a quantitative perspective, Shadish et al. (2002) emphasized how methodological strategies contribute to study rigor through credibility and integrity. This focus aligns with Sandelowski’s (1993) and Tracy’s (2010) view that qualitative rigor is concerned with overall methodological fidelity. A formative approach to **quality** focuses on credibility and transparency, highlighting trustworthiness, transferability, dependability, and

confirmability. Actions for establishing quality involve precisely documenting the research process, describing the study, and exploring alternative approaches to strengthen conclusions. Shadish et al. (2002) linked quantitative quality to a validity typology, focused on demonstrating the “correctness” of inferences. Shenton (2004) and Tracy (2010), following Lincoln and Guba (1985), defined quality in qualitative research as establishing trustworthiness to demonstrate credibility.

Recommending Four Common Evaluative Research Practices

Four cross-paradigmatic practices—visual models, instrument pilots, analysis pilots, and triangulation—are recommended to formatively enhance rigor and quality. This section defines each practice and suggests implementation timing and documentation.

Visual Model

A visual model is a diagram that illustrates and monitors study procedures to ensure alignment among research aims, study design, and methodology. It is implemented during the conceptual and planning phases and maintained through the interpretation of findings and results. Barker and Pistrang (2005) defined alignment as pairing the explanation of context and purpose with the use of appropriate methods, which can be illustrated through visual modeling. Often called research or design diagrams, visual models are commonly used in mixed methods studies to display decisions related to the problem statement, study purpose, research questions, theoretical framework(s), data collection procedures, and analysis (Tashakkori et al., 2020). Shenton (2004) referred to visual models as diagrammatic audit trails, which may take a “data-oriented” (traditional) or theoretical (holistic) approach to “trace the course

of the research step-by-step via decisions made and procedures described” (p. 72), ensuring alignment and promoting confirmability.

Moher et al. (2010) proposed a theoretically informed visual model to illustrate participant flow in randomized trials, aligning with CONSORT guidelines and emphasizing internal and external validity. The authors stressed validity should be established during planning, maintained throughout the trial, and reported accurately. This method can also be adapted to other quantitative studies to track data flow and ensure transparency. Creamer (2021) recommended diagramming analysis procedures in qualitative research, especially for complex studies with diverse data sources. Leech and Onwuegbuzie (2010) referred to this process as “outlining” in mixed methods, linking planning and implementation to diagramming with process models (see Johnson & Onwuegbuzie, 2004).

Monitoring and documenting the inquiry process prioritizes transparency and systematicity (Morse et al., 2002) while emphasizing data collection alignment and design consistency. This approach enables researchers to review decisions during the inquiry process and maintains documentation of decisions (Creamer, 2021). As both an internal practice and external product, a visual model can help monitor consistency and accuracy, complementing an audit trail to enhance external verification (Creswell & Miller, 2000).

Instrument Pilot

An **instrument pilot** evaluates data collection tools to ensure they accurately and reliably reflect the concepts studied. It is implemented during a pilot data collection phase to ensure the instruments are accurate, aligned, and representative. In survey-

1
2
3
4
5
6
7
8 based research, this involves assessing how well an instrument measures a concept
9 (Fowler, 2014). In qualitative studies, where the researcher is the primary instrument,
10 preliminary interviews or observations are used to assess the accuracy of procedures
11 (Patton, 2014). Malmqvist et al. (2019) emphasized that using consistent, well-tested
12 instruments increases confidence in data trustworthiness. This aligns with Fowler's
13 quantitative (2014) and Patton's (2014) evaluative focus on internal efficacy and
14 external extrapolation, as well as Lincoln and Guba's (1985) and Shenton's (2004)
15 parallel forms of ensuring validity. An instrument pilot can demonstrate several forms of
16 validity.
17

- 18 1. Content validity relies on experts to evaluate whether the instrument accurately
19 represents the intended concepts or theoretical-based indicators (Fowler, 2014).
20 This is sometimes referred to as expert review in qualitative research (Patton,
21 2014). Face validity similarly relies on non-content experts for validation (Denzin,
22 2017).
23
- 24 2. Theoretical validity evaluates both descriptive (concept accuracy) and
25 interpretive validity (concept relationships) to ensure alignment with the research
26 problem. This may involve testing contrasting interpretations with alternative
27 theories, comparing similar instruments to assess convergent and discriminant
28 validity, or assessing consistency across measures (Malmqvist et al., 2019;
29 Shadish et al., 2002).
30
- 31 3. Criterion validity compares individual components of an instrument to assess
32 alignment with a "gold standard" measure (Shadish et al., 2002). A comparable
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

method for piloting context- and content-specific qualitative instruments involves asking several questions in the same way within a similar setting (Patton, 2014).

4. Equity-centered validity considers the research context, including the history of the location, the program, and the people, and how power is distributed in a community (Waterman et al., 2022). This may involve inviting target population members to review and audit questions, prompts, or other forms for cultural sensitivity. Centering equity as a validity proposition helps identify outliers that could otherwise mask divergent perspectives or perpetuate marginalization (Mbah et al., 2024).

Analysis Pilot

An **analysis pilot** evaluates pilot data to assess whether the evidence accurately represents the intended concepts. It is implemented during a pilot data collection phase to establish credibility and ensure that the findings reliably yield the intended results. This practice helps establish measurement validity by collecting and verifying data throughout the study, ensuring that the constructs are accurately represented. For example, survey responses that are too similar might indicate an inadequate measure or the need to broaden the scale (Fowler, 2014). Likewise, interview pilot results might uncover further topics that need to be included in an updated protocol.

In Gadella Kamstra's (2022) study of teacher motivation, unexpected statistical results during the analysis pilot revealed the need for a "methodological transformation." The questionnaire, designed to identify and understand teacher (de)motivation, did not allow participants to fully express the struggles they were experiencing. This raised concerns about the questionnaire's appropriateness, as it only captured a superficial

1
2
3
4
5
6
7
8 understanding. These findings prompted the development of interview prompts—an
9 adaptive data collection instrument based on reliable items from the questionnaire.

10
11
12 During an analysis pilot, qualitative researchers may use inter-rater or inter-coder
13 comparisons to enhance interpretation. This involves asking multiple observers to
14 collect or review data and compare interpretations (O'Connor & Joffe, 2020). In direct
15 interaction research, a valuable analysis pilot technique is member checking, where
16 participants are asked to verify the authenticity or accuracy of the data and/or
17 conclusions (Lincoln & Guba, 1985). Morse (2015) recommended integrating member
18 checking throughout a study, as summative checks can be problematic if participants
19 disagree with interpretations, especially when participants may not recognize their own
20 perspectives.
21
22
23
24
25
26
27
28

29 **Triangulation**

30
31 Triangulation uses an additional inquiry approach to explore alternative or more
32 comprehensive understandings of a research problem. It is implemented during data
33 collection and analysis and complements or verifies findings to strengthen research
34 quality, credibility, confirmability, and applicability. Triangulation can involve multiple
35 methods, investigators, theories, and/or data sources. It differs from mixed methods
36 research¹, which integrates diverse approaches to deepen understanding (Tashakkori
37 et al., 2020). Evaluative triangulation is linked to confirmation and corroboration (Greene
38 et al., 1989) and parallels cross-validation (Maxwell, 2013; Patton, 2014). Mathison
39
40
41
42
43
44
45
46
47

48 ¹ Contemporary mixed methods scholars argue that triangulation contradicts broader goals of exploration
49 and understanding subjective meanings (Fetters & Molina-Azorin, 2017). While Greene et al. (1989)
50 viewed triangulation as a "constrained-narrow" design. Bryman (2002) critiqued triangulation as a naively
51 realist approach focused on achieving convergent validity. This perspective aligns with Flick's (2017) view
52 that a strong triangulation program should generate new knowledge, rather than merely confirming
53 existing findings.
54
55
56
57
58
59
60

(1988) suggested triangulated approaches provide a rich and complex picture of evidence by embedding data within a broader context and knowledge. This perspective aligns with Maxwell's (2004) emphasis on the role of meaning and interpretive understanding in causal explanations, reinforcing the notion that triangulation is not merely about confirmation but also about providing deeper insights and transferable inferences.

In a review of research on Gay-Straight Alliances (GSA), Poteat et al. (2017) shared how they triangulated multiple data sources. Recommendations included supplementing youth-reported data with advisor-reported data and records, adding data from additional stakeholders to provide a more comprehensive view of how GSAs and similar groups operate within a broader context, and adding contextual data about schools and districts within the domain of operation such as a school performance and district policies. Using multiple data sources helped account for variability in activities and experiences.

Four inquiry approaches are suggested as formative practices to improve quality and credibility across methodologies, drawing on Denzin's (2017) recommendations.

1. Method triangulation relies on multiple methods for data collection. An example of within-method triangulation is combining survey responses with existing datasets or pairing interviews with observations. In contrast, between-method triangulation might involve integrating survey responses with interview data (Shenton, 2004).
2. Investigator triangulation involves two or more researchers independently collecting data or conducting analyses. Flick (2017) recommended including investigators with diverse theoretical, epistemological, or methodological

perspectives. Involving additional researchers can help to address limitations, improve inference quality, and explore tensions in study design and analysis (Archibald, 2016).

3. Theory triangulation uses multiple theories to support data or address plausible rival explanations (Barker & Pistrang, 2005). Flick (2017) noted this often co-occurs with investigator triangulation. Using multiple theories can strengthen confidence in findings by testing them against alternative hypotheses. However, theories must be clearly defined and justified to align with the phenomenon being studied.
4. Data triangulation uses multiple data sources to verify findings. Shenton (2004) noted it enables cross-checking data, leading to a deeper understanding of the findings. Incorporating diverse sources, such as different informants or observations at various times, enhances the credibility and depth of interpretations (Patton, 2014) and can be effective for confirming or challenging conclusions.

Greene et al. (1989) argued that different inquiry forms should be framed within a unified paradigmatic approach. However, Flick (2017) cautioned against overfocusing on validating a specific paradigm, stressing that triangulation should center on the research problem. Additionally, convergence, inconsistency, and contradictions of triangulated points should be described when documenting and reporting interpretations (Mathison, 1988).

Final Considerations

This article recommended four practices for monitoring and evaluating social science research across methodologies. These practices are not intended to replace method-specific techniques. Instead, they introduce formative checks (Black & Wiliam, 2009), often overlooked in research (Morse et al., 2002), to facilitate ongoing evaluation throughout the process (Dellinger & Leech, 2007; Maxwell, 2013). Consistent use and documentation are essential to ensure representation, inclusion, and credibility in research (Waterman et al., 2022). By offering common approaches, the article provides guidelines for researchers to monitor and demonstrate rigor and quality in their work.

References

- Archibald, M. M. (2016). Investigator triangulation: A collaborative strategy with potential for mixed methods research. *Journal of Mixed Methods Research*, 10(3), 228–250.
- Barker, C., & Pistrang, N. (2005). Quality criteria under methodological pluralism: Implications for conducting and evaluating research. *American Journal of Community Psychology*, 35, 201-212.
- Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21, 5-31 (2009).
- Cook, T. D. (2009). "Waiting for life to arrive": A history of the regression-discontinuity design in psychology, statistics, and economics. *Journal of Econometrics*, 142(2), 636–654.
- Creamer, E. G. (2021). Leveraging an integrated visual display for case-based analysis in mixed method research. *International Journal of Qualitative Methods*, 20.

- 1
2
3
4
5
6
7
8
9 Creswell, J. W., & Miller, D. L. (2000). Determining validity in qualitative inquiry. *Theory*
10 *into Practice*, 39(3), 124-130.
- 11
12 Dellinger, A. B., & Leech, N. L. (2007). Toward a unified validation framework in mixed
13 methods research. *Journal of Mixed Methods Research*, 1(4), 309-332.
- 14
15 Denzin, N. K. (2017). *The research act: A theoretical introduction to sociological*
16 *methods*. Transaction publishers.
- 17
18 Flick, U. (2017). Mantras and myths: The disenchantment of mixed-methods research
19 and revisiting triangulation as a perspective. *Qualitative inquiry*, 23(1), 46-57.
- 20
21 Fowler, F. J. (2014). *Survey research methods* (5th ed.). Sage.
- 22
23 Gadella Kamstra, L. S. (2022). Transforming a methodological dilemma into a rewarding
24 research opportunity. *International Journal of Social Research*
25 *Methodology*, 25(6), 741-755.
- 26
27 Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework
28 for mixed-method evaluation designs. *Educational Evaluation and Policy*
29 *Analysis*, 11(3), 255-274.
- 30
31 Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research
32 paradigm whose time has come. *Educational Researcher*, 33(7), 14-26.
- 33
34 Leech, N. L., & Onwuegbuzie, A. J. (2010). Guidelines for conducting and reporting
35 mixed research in the field of counseling and beyond. *Journal of Counseling &*
36 *Development*, 88(1), 61-69.
- 37
38 Lincoln, Y. S., & Guba, E. G. (1985). *Effective evaluation: Improving the usefulness of*
39 *evaluation results through responsive and naturalistic approaches*. Jossey-Bass.
- 40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3
4
5
6
7
8
9 Malmqvist, J., Hellberg, K., Möllås, G., Rose, R., & Shevlin, M. (2019). Conducting the
10 pilot study: A neglected part of the research process? Methodological findings
11 supporting the importance of piloting in qualitative research studies. *International*
12 *Journal of Qualitative Methods*, 18, 1-11.
13
14
15
16 Mathison, S. (1988). Why triangulate? *Educational Researcher*, 17(2), 13-17.
17
18 Maxwell, J. A. (2004). Causal explanation, qualitative research, and scientific inquiry in
19 education. *Educational Researcher*, 33(2), 3-11.
20
21 Maxwell, J. A. (2013). *Qualitative research design: An interactive approach* (3rd ed.).
22 SAGE Publications.
23
24
25 Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P. J., ...
26 & Altman, D. G. (2010). CONSORT 2010 explanation and elaboration: updated
27 guidelines for reporting parallel group randomised trials. *BJM*, 1-28.
28
29
30 Morse, J. M., Barrett, M., Mayan, M., Olson, K., & Spiers, J. (2002). Verification
31 strategies for establishing reliability and validity in qualitative research.
32 *International Journal of Qualitative Methods*, 1(2), 13-22.
33
34
35 Morse J. M. (2015). Critical analysis of strategies for determining rigor in qualitative
36 inquiry. *Qualitative Health Research*, 25(9), 1212-1222.
37
38
39 O'Connor, C., & Joffe, H. (2020). Intercoder reliability in qualitative research: debates
40 and practical guidelines. *International Journal of Qualitative Methods*, 19,
41 1609406919899220.
42
43
44
45 Onwuegbuzie, A. J., & Leech, N. L. (2005). On becoming a pragmatic researcher: The
46 importance of combining quantitative and qualitative research
47
48
49
50
51
52
53
54
55
56
57
58
59
60

methodologies. *International Journal of Social Research Methodology*, 8(5), 375-387.

Ong, B. K. (2024). Ways of establishing rigour in the Abductive Research Strategy (ARS). *International Journal of Social Research Methodology*, 27(6), 707-719.

Patton, M. Q. (2014). *Qualitative research & evaluation methods: Integrating theory and practice*. Sage.

Poteat, V. P., Yoshikawa, H., Calzo, J. P., Russell, S. T., & Horn, S. (2017). Gay-straight alliances as settings for youth inclusion and development: Future conceptual and methodological directions for research on these and other student groups in schools. *Educational Researcher*, 46(9), 508-516.

Sandelowski, M. (1993). Rigor or rigor mortis: The problem of rigor in qualitative research revisited. *Advances in Nursing Science*, 16(2), 1-8.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton, Mifflin and Company.

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information*, 22(2), 63-75.

Tashakkori, A., Johnson, R. B., & Teddlie, C. (2020). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Sage.

Tracy, S. J. (2010). Qualitative quality: Eight "big-tent" criteria for excellent qualitative research. *Qualitative inquiry*, 16(10), 837-851.

1
2
3
4
5
6
7
8 Van Assche, J., et al. (2021). A meta-analysis on methodological choices in social
9 psychology. *Psychological Methods*, 26(1), 67–85.

10
11
12 Waterman, S. J., Baker, G. R., Henning, G. W., & Lundquist, A. E. (2023). Equity-
13 centered assessment: Varying approaches and lenses. In G. Henning et al
14 (Eds.), *Reframing Assessment to Center Equity* (pp. 112-128). Routledge.
15
16

17 **Acknowledgement**

18
19 ChatGPT (version 4) was used for language improvement, in accordance with Taylor &
20 Francis policies. The tool assisted in refining the clarity and flow of the manuscript.
21
22

23 ChatGPT was not used for text generation without rigorous review and revision.
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60