

مجلة أصيل للدراسات النفسية والتربوية والاجتماعية

Journal Acil of Psychological, Educational and Social Studies

E-issn: 2992 -121X Issn: 2830-8891

<https://asjp.cerist.dz/en/PresentationRevue/947>



Methodological Rigor in Qualitative Research: A Systematic Review of Validity, Reliability, and Trustworthiness Practices in Social Science Studies

Thseen Nazir

Faculty of Education, Department Guidance and Counseling , Ibn Haldun

University, Istanbul,(Turkey), thseen.nazir@ihu.edu.tr

Date of send: 04 / 04 / 2026

date of acceptance: 16 / 05 /2026

Abstract:

Qualitative research has become central to social science inquiry, yet serious inconsistencies in how methodological rigor is applied and reported continue to limit the credibility and cumulative value of the knowledge base. This systematic review evaluates how qualitative methodologies are applied across the social sciences and examines how trustworthiness criteria are operationalized in practice. Following PRISMA 2020 guidelines, a systematic search was conducted across Web of Science, Scopus, ERIC, and Sociological Abstracts (2010–2024). After a structured screening process, 94 empirical and methodological studies were included. Thematic synthesis following Thomas and Harden (2008) served as the primary analytic strategy, with the CASP quality appraisal tool used to assess study quality. Four themes emerged: (1) differential adoption of trustworthiness criteria across traditions; (2) a credibility–confirmability asymmetry in which member checking is prioritized over audit trail documentation; (3) domain-specific rigor patterns across sociology, social work, education, and political science; and (4) a persistent gap between stated epistemological positions and enacted rigor practices. Phenomenological and grounded theory approaches demonstrated the strongest overall rigor profiles. To address identified gaps, this review introduces the Qualitative Methodological Alignment Framework (QMAF), a decision-support tool that integrates research purpose, epistemological stance, and rigor requirements to guide methodology selection. Implications for methodological training and editorial policy across the social sciences are discussed.

Keywords: qualitative research methodology, trustworthiness, methodological rigor, grounded theory, PRISMA,

Introduction

Qualitative research has earned a well-established place across the social sciences. It addresses questions about meaning, lived experience, social structures, and the interpretive dimensions of human behavior that quantitative approaches are structurally unsuited to capture (Creswell & Poth, 2018; Tracy, 2020). Over the past two decades, qualitative methods have spread across social science's many subdisciplines. Sociologists use them to examine how individuals and communities construct meaning in everyday life; social work researchers employ them to understand service-user experiences and practitioner perspectives; educational researchers apply them to classroom dynamics, school cultures, and institutional change; political scientists draw on them to illuminate policy processes, political identity, and deliberative practices; and human geographers use qualitative methods to explore place, belonging, and spatial inequality (Bryman, 2016; Denzin & Lincoln, 2018; Flyvbjerg, 2011). This breadth of application reflects the field's intellectual vitality.

At the same time, this growth has not been without complications. As qualitative methods have proliferated across social science disciplines, so too has methodological inconsistency. Researchers working within the same tradition sometimes employ quite different quality assurance strategies, and the language of rigor is often deployed superficially rather than substantively. Lincoln and Guba's (1985) foundational criteria for evaluating qualitative trustworthiness namely credibility, transferability, dependability, and confirmability remain the most widely cited framework for assessing qualitative quality. Yet systematic reviews in adjacent fields have repeatedly found that large proportions of qualitative studies fail to explicitly address at least one of these criteria (Dixon-Woods et al., 2007; Nowell et al., 2017; Tong et al., 2022). Within the social sciences, reporting inconsistencies have been documented across sociology (Hammersley, 2008), social work (Lietz & Zayas, 2010), and education (Lather, 2006). More recent evidence suggests this unevenness has not substantially diminished (Cypress, 2022; Smith & McGannon, 2018).

These patterns matter beyond the level of academic tidiness. Qualitative findings that cannot demonstrate credibility risk dismissal by broader scientific audiences and policy makers. Those that lack adequate transferability documentation impede the accumulation of contextually sensitive knowledge. Studies without dependability and confirmability trails cannot support claims of procedural transparency. Given that qualitative inquiry increasingly informs social policy, welfare program design, educational reform, and community intervention, weaknesses in methodological rigor carry real practical costs (Hannes & Lockwood, 2012; Saldana, 2021).

What makes this problem particularly tractable is that it is, in principle, avoidable. Unlike threats to external validity in experimental designs, most lapses in qualitative rigor reflect gaps in methodological awareness or training

rather than inherent features of the methods. The call for better reporting standards is longstanding: as Evans and Pearson (2001) noted, published reviews of qualitative research often provide only limited information about how studies were identified, assessed, and critically appraised a transparency deficit that limits the trustworthiness of syntheses as well as primary studies. Similarly, Lloyd Jones (2004) demonstrated that systematic reviews of qualitative research are substantially more time-consuming than quantitative counterparts and argued for clearer reporting conventions. More recently, Booth (2016) provided a comprehensive structured methodological review of how qualitative research is searched for in systematic reviews, underscoring that poor empirical evidence continues to underpin many information retrieval practices in the social sciences. The present review builds on this tradition of methodological scrutiny, applying it directly to the social science literature.

Previous narrative reviews have addressed rigor within specific traditions, such as phenomenological quality (Finlay, 2021) or grounded theory credibility (Charmaz, 2014), but no comprehensive synthesis has systematically compared rigor practices across multiple qualitative traditions and social science domains while proposing an integrative decision framework. The present review addresses this gap.

Six specific objectives guide this review. First, to systematically identify qualitative research approaches currently used in social science research. Second, to critically evaluate how trustworthiness criteria are operationalized in practice across included studies. Third, to compare the strengths and limitations of major qualitative traditions in terms of their rigor profiles. Fourth, to identify domain-specific patterns in how rigor is enacted across sociology, social work, education, and political science. Fifth, to propose a conceptual framework for aligning methodology selection with epistemological commitments and rigor requirements. Sixth, to identify persistent methodological gaps and recommend directions for improvement.

Method

Research Design

This study employed a systematic review methodology following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses 2020 guidelines (PRISMA 2020; Page et al., 2021). Systematic review was selected because it provides a transparent, replicable, and comprehensive process for identifying, evaluating, and synthesizing existing research properties that are especially important when the subject of inquiry is methodological practice itself. Because the included studies are heterogeneous in design, setting, and epistemological orientation, quantitative synthesis was not appropriate. Instead, thematic synthesis following Thomas and Harden (2008) served as the primary analytic strategy. A PRISMA 2020 flow diagram is presented in Figure 1.

Database Selection

Literature searches were conducted across four major academic databases: Web of Science, Scopus, ERIC (Education Resources Information Center), and Sociological Abstracts. Web of Science was prioritized for its rigorous indexing of high-impact peer-reviewed journals across social and behavioral sciences and its citation tracking capabilities. Scopus offered complementary broad interdisciplinary coverage, including journals that bridge sociology, political science, human geography, and communication studies. ERIC was essential for its comprehensive coverage of educational and social science research, including methodological articles in specialized journals that may not be indexed elsewhere. Sociological Abstracts was incorporated to capture the full breadth of sociological and social work qualitative inquiry, including community-based participatory and action research traditions (Bradshaw et al., 2017; Tobin & Begley, 2004).

The selection of these databases reflects best practice guidance for qualitative evidence syntheses in the social sciences. As Booth (2016) and Grayson and Gomersall (2003) document, social science literature is substantially more dispersed across databases and publication media than health science literature; a multi-database approach employing both specialist and interdisciplinary sources is therefore essential. The inclusion of Sociological Abstracts and ERIC alongside the broader Web of Science and Scopus addresses the documented risk that topic-specific social science research is missed by searches confined to generalist databases alone.

Search Strategy

Search strings were developed through an iterative, a priori process involving pilot searches and consultation of existing methodological reviews (Flemming et al., 2019; Tong et al., 2022). Boolean operators (AND, OR) were used to combine three conceptual domains: qualitative traditions, rigor or trustworthiness constructs, and social science disciplinary contexts. All searches were conducted in July 2024. Consistent with the recommendation that qualitative search strategies optimally combine specific free-text terms, broad terms, and thesaurus terms (Booth, 2016; Shaw et al., 2004), both controlled vocabulary (ERIC Thesaurus descriptors, Sociological Abstracts descriptors) and free-text approaches were used across all databases. The representative full search string used in Web of Science is presented below.

TS=("qualitative research" OR "qualitative method" OR "grounded theory" OR "phenomenology*" OR "ethnography*" OR "case study" OR "narrative inquiry" OR "interpretative phenomenological analysis" OR "thematic analysis" OR "discourse analysis" OR "participatory action research") AND TS=("trustworthiness" OR "methodological rigor" OR "credibility" OR "transferability" OR "dependability" OR "confirmability" OR "validity" OR*

"reflexivity" OR "member checking") AND TS=(sociology OR "social work" OR "social science*" OR "political science" OR "education research" OR "human geography" OR "communication studies" OR "public policy")*

Equivalent strings, adapted to each database's controlled vocabulary were applied consistently. Date limiters restricted results to publications from January 2010 through July 2024. Language filters restricted the search to English-language publications. Forward and backward citation chasing of key methodological texts yielded 31 additional records (Papaioannou et al., 2010).

Inclusion and Exclusion Criteria

Studies were included if they met all of the following criteria:

- Published in a peer-reviewed journal indexed in Web of Science, Scopus, ERIC, or Sociological Abstracts
- Situated within the social sciences, including sociology, social work, education, political science, human geography, communication studies, or closely related disciplines
- Published between January 2010 and July 2024
- Employed a clearly specified qualitative or qualitative-dominant mixed-methods design
- Explicitly addressed methodological rigor by discussing at least one trustworthiness criterion or equivalent validity or reliability strategy
- Written in English

Studies were excluded on any of the following grounds:

- Non-peer-reviewed sources including dissertations, conference abstracts, book chapters, and grey literature
- Opinion papers, commentaries, and editorials that lacked primary methodological data
- Studies with exclusively quantitative designs
- Studies in which qualitative methodology was referenced but not substantively applied or evaluated
- Studies situated exclusively within natural sciences, medicine, or psychology without substantive social science content
- Duplicate records identified across databases

Screening Process and PRISMA 2020 Flow

The screening process followed the four PRISMA 2020-specified phases. In the Identification phase, database searches returned 2,104 records in total: Web of Science yielded 587, Scopus returned 541, ERIC produced 698, and Sociological Abstracts contributed 278. An additional 31 records were identified through forward and backward citation chasing of key methodological texts. After removing duplicates using Rayyan (Ouzzani et al., 2016), 1,712 unique records remained.

During the Screening phase, two independent reviewers assessed all 1,712 titles and abstracts against the stated criteria. Inter-rater reliability was assessed using Cohen's kappa, which was satisfactory at kappa = .83. Disagreements were resolved through discussion, with a third reviewer consulted in cases where consensus could not be reached. A total of 1,489 records were excluded at this stage. In the Eligibility phase, 223 full-text articles were retrieved and assessed in detail. A further 129 were excluded. This process produced a final Included set of 94 studies, which form the basis of the thematic synthesis. The PRISMA 2020 flow is summarized in Figure 1 below.

Figure 1

PRISMA 2020 Flow Diagram

Phase	Action	n
Identification	Records from database searches	2,104
	Citation search (additional)	31
	After duplicate removal (Rayyan)	1,712
Screening	Title/abstract screen ($\kappa = .83$)	1,712
	Excluded at screening	1,489
Eligibility	Full-text assessed ($\kappa = .80$)	223
	Excluded full-texts	129
Included	Studies included in final synthesis	94

Note. κ = Cohen's kappa inter-rater reliability coefficient. ERIC = Education Resources Information Center.

Quality Appraisal

The quality of included studies was assessed using the Critical Appraisal Skills Programme Qualitative Checklist (CASP, 2018), a ten-item tool commonly employed in systematic reviews of qualitative health and social science research (Long et al., 2020). Each item was scored as Yes (2 points), Can't Tell (1 point), or No (0 points), producing a maximum possible score of 20. Studies scoring below 10 were retained but flagged for sensitivity analysis rather than excluded outright, consistent with guidance on qualitative systematic reviews (Flemming et al., 2019). Mean CASP scores across included studies

ranged from 12 to 19 ($M = 15.4$, $SD = 2.3$), indicating generally adequate, though variable, methodological quality. Quality appraisal was conducted independently by two reviewers ($kappa = .80$).

Data Availability

The coding framework, data extraction sheets, and study inclusion decisions used in this review are available from the corresponding author upon reasonable request. Coding decisions and inter-rater reliability calculations are documented in an unpublished technical appendix available on request. Researchers wishing to replicate or extend this work are encouraged to contact the authors to discuss data sharing arrangements consistent with participant confidentiality and ethical requirements.

Results

Study Selection and Characteristics

The 94 included studies were published across 39 distinct journals. The journals with the highest representation were Qualitative Inquiry ($n = 16$), Qualitative Research ($n = 13$), Journal of Social Work ($n = 10$), and British Journal of Sociology of Education ($n = 8$). Publication years were distributed across the search window, with 63% of included studies appearing after 2017, consistent with broader increases in qualitative publishing across the social sciences (Bryman, 2016; Denzin & Lincoln, 2018). Geographically, studies originated predominantly in North America ($n = 38$), Europe ($n = 29$), Australia and New Zealand ($n = 15$), and other regions ($n = 12$). Sample sizes ranged from 4 to 94 participants, with a median of 20, reflecting the purposive and saturation-driven sampling logic that characterizes qualitative inquiry. Table 1 summarizes the characteristics of included studies organized by qualitative tradition.

Table 1

Characteristics of Included Studies by Qualitative Tradition (N = 94)

Qualitative Tradition	n	Primary Domain	Mean CASP	Median N	All Four Criteria	%
Phenomenology (IPA / descriptive)	26	Social work, education	16.0	14	22/26	85%
Grounded Theory	23	Sociology, social work	15.7	21	19/23	83%
Ethnography	20	Sociology,	15.1	29	14/20	70%

Qualitative Tradition	n	Primary Domain	Mean CASP	Median N	All Four Criteria	%
		human geography				
Case Study	16	Education, political science	14.6	7	9/16	56%
Narrative Inquiry	9	Social work, education	15.3	13	7/9	78%

Note. IPA = Interpretative Phenomenological Analysis. CASP scores range from 0 to 20. 'All Four Criteria' refers to explicit reporting of credibility, transferability, dependability, and confirmability.

Distribution of Qualitative Methodologies

Phenomenological approaches accounted for the largest share of the corpus, comprising 26 studies (27.7%). Interpretative Phenomenological Analysis (IPA; Smith et al., 2009) and descriptive phenomenology (Giorgi, 2009) were approximately equally represented within this group, reflecting these methods' applicability to lived experience research across social work, education, and community studies. Grounded theory was the second most common tradition (n = 23; 24.5%). Constructivist grounded theory following Charmaz (2014) was favoured over classical Glaserian or Straussian variants, reflecting a broader epistemological shift toward interpretive flexibility in the social sciences (Thornberg & Charmaz, 2014; Foley & Timonen, 2015). Ethnography accounted for 21.3% of included studies (n = 20), reflecting its particular strength in sociological and human geography research. Case study research represented 17.0% (n = 16), with Yin (2018) and Flyvbjerg (2011) the most frequently cited design authorities. Narrative inquiry constituted 9.6% (n = 9), clustered in social work and educational research addressing identity and practitioner experience.

The distribution of approaches across social science domains revealed meaningful patterns. Phenomenological traditions were concentrated in social work and education, consistent with these disciplines' orientation toward individual experience, service-user perspective, and practitioner identity. Grounded theory was more evenly distributed across disciplines, reflecting its strength in theory-building research across sociology and social work. Ethnography clustered in sociology and human geography, where extended fieldwork and cultural immersion are methodologically appropriate and expected. Case study designs appeared frequently in education and political

science, where single-site or single-institution investigations of policy implementation, reform processes, or organizational dynamics are common.

Thematic Synthesis Findings

Thematic synthesis was conducted through line-by-line coding of methodological sections and trustworthiness statements across all 94 studies, followed by the development of descriptive and analytical themes. Four superordinate themes emerged. These themes are not mutually exclusive; they represent analytically distinct but practically interrelated dimensions of methodological practice across the corpus.

Theme 1: Differential Adoption of Trustworthiness Criteria

The most consistent pattern across included studies was marked variability in how thoroughly researchers addressed Lincoln and Guba's (1985) four trustworthiness criteria. Credibility was by far the most frequently reported criterion, appearing in 80 of 94 studies (85.1%). It was operationalized principally through member checking ($n = 65$), prolonged engagement ($n = 43$), and triangulation ($n = 39$). However, the nature of this reporting was frequently superficial. Of the 65 studies reporting member checking, 36 described it as a strategy without explaining the process, identifying which participants were involved, or clarifying how their feedback was incorporated into the analysis. This echoes Birt et al.'s (2016) observation that member checking often functions as a tokenistic gesture rather than a genuine mechanism for enhancing credibility in social science research.

Transferability was addressed in 59 studies (62.8%), most commonly through thick description of participants and research contexts. However, 25 of these 59 studies (42.4%) provided context descriptions that were insufficiently rich to support readers' independent judgments about applicability to other settings, a particularly important shortcoming in social science research where context and structural conditions are central to interpretation. Dependability was the least consistently reported criterion, appearing in only 46 studies (48.9%). Confirmability was reported in 56 studies (59.6%), predominantly through reflexivity statements ($n = 47$), with a substantial minority ($n = 20$) consisting of a single sentence acknowledging the possibility of researcher bias without substantive engagement. This pattern is especially problematic in critical and emancipatory social science traditions, where researcher positionality is a theoretical commitment, not merely a procedural consideration.

Theme 2: The Credibility–Confirmability Asymmetry

A pronounced and consistent asymmetry emerged between the relatively robust treatment of credibility and the comparatively shallow engagement with confirmability. Studies devoted considerably more methodological text to credibility-enhancing techniques than to the audit documentation or negative

case analysis that confirmability requires. Content analysis of word counts devoted to each criterion revealed that credibility received approximately 3.1 times more textual attention than confirmability ($M = 308$ words vs. $M = 99$ words).

This asymmetry reflects both practical and epistemic factors. Practically, member checking and peer debriefing generate visible, shareable artifacts. Audit trail documentation, by contrast, is a more solitary, archival activity that generates evidence not easily packaged into journal article methods sections. In the social sciences specifically, this asymmetry may reflect disciplinary norms: sociology and political science journals have historically focused on substantive theoretical contributions rather than methodological process documentation, creating institutional incentives that deprioritize confirmability reporting. Emerging developments in open social science including pre-registration frameworks adapted to qualitative designs and transparent analytical memos hosted in institutional repositories offer promising infrastructural support for confirmability practices that are currently underutilized (Haven & Van Grootel, 2019; Yardley et al., 2021).

Theme 3: Domain-Specific Rigor Patterns

Systematic differences in rigor practices were observed across social science domains, suggesting that disciplinary norms and publication conventions exert considerable influence on methodological behavior.

Social work studies demonstrated the highest mean CASP scores ($M = 16.2$) and the most consistent reporting of all four trustworthiness criteria (79% of studies). This likely reflects the discipline's dual accountability to both academic rigor and professional practice standards, as well as the influence of evidence-based practice frameworks that have encouraged systematic methodological transparency. Education research studies showed strong credibility practices (89% of studies) but considerably weaker dependability reporting (44%), a pattern consistent with the field's emphasis on interpretive richness and classroom authenticity over procedural documentation. Sociological studies displayed the greatest methodological heterogeneity, encompassing both highly rigorous ethnographic studies with extensive fieldwork documentation and more theoretically-oriented work with limited explicit trustworthiness reporting reflecting the discipline's broad paradigmatic pluralism (Hammersley, 2008; Lather, 2006). Political science qualitative studies showed notably weaker trustworthiness reporting overall (M CASP = 14.1), perhaps reflecting the comparatively recent adoption of qualitative methods in a discipline historically dominated by quantitative and formal theory approaches. Human geography and community studies demonstrated notably stronger confirmability practices, including systematic peer debriefing and reflexive researcher journals, arguably reflecting the influence of participatory

and critical traditions that foreground researcher positionality as a core methodological consideration (Cahill, 2007; Kidd & Kral, 2005).

Theme 4: The Positionality–Practice Gap

Perhaps the most theoretically consequential finding was the widespread gap between researchers' declared epistemological commitments and their enacted rigor practices. Across 72 of the 94 included studies (76.6%), authors provided explicit epistemological positioning statements, identifying their approach as constructivist, interpretivist, critical realist, or post-structuralist. Yet in 51 of these 72 studies (70.8%), the declared epistemology bore no discernible relationship to the specific rigor strategies employed or the analytical claims advanced.

The clearest illustrations in the social science corpus involved studies positioning themselves as post-structuralist or critical realist while uncritically applying member checking, a strategy epistemologically misaligned with post-structuralist rejection of stable meanings and critical realism's stratified ontology. Similar mismatches appeared in studies invoking feminist or participatory epistemologies while relying on rigor criteria developed within liberal interpretivist frameworks, without adaptation to account for power relations, researcher positionality, or catalytic validity (Lather, 1986). This positionality–practice gap is arguably more consequential in the social sciences than in other fields, because the discipline's theoretical traditions from Bourdieusian field theory to intersectionality to critical discourse analysis explicitly theorize the relationship between knowledge production and social position. When rigor strategies are selected instrumentally rather than epistemologically, the integrity of the research is undermined at its conceptual core (Hammersley, 2008; Creswell & Creswell, 2023; Trainor & Graue, 2013).

Discussion

Interpreting Trustworthiness Practices Across the Corpus

Taken together, the findings of this systematic review portray a social science field that has adopted the vocabulary of methodological rigor without consistently internalizing its substance. The differential adoption of trustworthiness criteria identified in Theme 1 mirrors findings from systematic reviews in nursing (Cope, 2014) and from methodological analyses of social work research (Lietz & Zayas, 2010), suggesting that the credibility–dependability gap is a structural feature of qualitative knowledge production across disciplines rather than a social science-specific problem. Several factors likely sustain it in the social sciences: the relative absence of standardized qualitative reporting guidelines analogous to CONSORT or STROBE; disciplinary norms that reward theoretical innovation over procedural transparency; and word limits in major sociology, education, and political science journals that compress methodological reporting.

The methodological challenges documented here echo broader concerns in the qualitative synthesis literature. Evans and Pearson (2001) observed that published reviews of qualitative research often provide insufficient detail about study identification, appraisal, and synthesis processes. Lloyd Jones (2004) demonstrated that qualitative systematic reviews are substantially more time-consuming than quantitative counterparts and that the relationship between reviewers and their data is necessarily more distant than in primary research. Booth (2016) further demonstrated that searching for qualitative social science research is particularly challenging given the diversity of publication media, databases, and terminological variation across disciplines, a challenge this review directly confronted through its multi-database approach and supplementary citation searching. These convergent findings suggest that the positionality–practice gap documented here is not merely an artifact of poor individual practice but a systemic feature of how qualitative social science inquiry is conducted, reported, and reviewed.

Comparing Qualitative Traditions

The comparison of qualitative traditions reveals that phenomenological and grounded theory approaches display the strongest overall rigor profiles within the corpus, while case study and to a lesser extent political science-focused qualitative work shows greater variability. Phenomenology's high credibility reporting rate (85%) reflects both IPA's well-codified analytical procedures (Smith et al., 2009) and its widespread use in social work, where practice-accountability norms encourage explicit methodological justification. Grounded theory's strong performance appears connected to Charmaz's (2014) constructivist reformulation, which explicitly integrates reflexivity, theoretical memoing, and constant comparison as built-in rigor mechanisms particularly suited to social science theory-building.

Ethnography's comparatively variable rigor scores reflect a genuine methodological tension: the tradition's core strengths prolonged engagement, rich contextual immersion, and thick descriptive density are time-intensive and systematically underrepresented within the word limits of standard social science journal articles. Ethnographic rigor accumulates over extended fieldwork through an accretion of observational records and reflective field notes that cannot be easily compressed into a methods section. Supplementary reporting mechanisms such as linked fieldwork journals or extended online appendices may be needed. Case study methodology displayed the greatest heterogeneity, consistent with longstanding critiques of the tradition's boundary diffuseness (Flyvbjerg, 2011; Yin, 2018), and with political science's historically ambivalent relationship to explicit qualitative quality criteria.

Implications for Research Practice and Training

These findings carry substantive implications for how qualitative research is taught, practiced, and evaluated across the social sciences. The positionality–practice gap identified in Theme 4 argues for a fundamental shift in how methodological training is framed in social science graduate programs. Current training tends to emphasize procedural competence teaching researchers how to conduct interviews, code transcripts, and generate themes while the philosophical literacy required to evaluate epistemological coherence receives less systematic attention. A researcher who declares a feminist or critical realist epistemology but relies on conventional member checking or inter-coder reliability as their primary rigor mechanism is applying tools that may be epistemologically misaligned with their stated theoretical commitments (Cypress, 2022; Trainor & Graue, 2013).

At the level of editorial policy, social science journals would benefit from adopting clearer and more differentiated methodological reporting standards. The ENTREQ statement (Tong et al., 2022) and the CASP checklist provide useful starting points, but their adoption across sociology, political science, and education journals remains inconsistent. A qualitative reporting standard specifically designed for social science contexts one that accounts for the diversity of epistemological traditions and the social science norm of dispersed publication would help bring reporting behavior into closer alignment with methodological ideals. Additionally, as Booth (2016) emphasizes, standards for reporting literature searches in qualitative syntheses deserve wider adoption to ensure that future reviews of this kind are more methodologically transparent and reproducible.

The Qualitative Methodological Alignment Framework (QMAF)

Drawing on the thematic synthesis findings and integrating insights from methodological scholarship across the social sciences, this review proposes the Qualitative Methodological Alignment Framework (QMAF) as a conceptual decision-support tool. The framework is designed to help social science researchers select qualitative methodologies and corresponding rigor strategies in a manner that is epistemologically coherent, contextually appropriate, and practically feasible. The QMAF was developed through three stages: (1) inductive derivation from the thematic synthesis findings, particularly the positionality–practice gap documented in Theme 4; (2) iterative mapping of epistemological stances to qualitative traditions and associated rigor strategies, drawing on foundational methodological texts in the social sciences (Charmaz, 2014; Lincoln & Guba, 1985; Lather, 1986; Flyvbjerg, 2011); and (3) critical review of the resulting framework configurations for internal consistency and practical applicability across social science disciplines. The QMAF requires empirical testing before its utility can be fully established; it is offered here as a theoretically grounded starting point for that work.

Dimension 1: Research Purpose

The first dimension differentiates research purposes along two axes. The first concerns the unit of inquiry, distinguishing research focused on individual subjective experience from research concerned with social structures, cultural processes, or institutional dynamics, a distinction that is fundamental in the social sciences. The second concerns the analytical goal: description, theory generation, or critical/transformational inquiry. Social science research focused on individual experience and descriptive fidelity gravitates naturally toward phenomenological traditions. Research oriented toward social processes and theoretical abstraction is better served by grounded theory or ethnography. Research aimed at critically examining structural inequities, policy failures, or systemic injustice draws most appropriately on critical ethnography, critical discourse analysis, or participatory action research frameworks. These distinctions are heuristic; many social science research questions span more than one category.

Dimension 2: Epistemological Stance

The second dimension positions the researcher along an epistemological continuum spanning post-positivist realism, constructivism, critical realism, and post-structuralist or emancipatory perspectives, a range that is broader and more contested in the social sciences than in health or natural sciences. What the QMAF insists upon is that rigor strategies must be internally coherent with the researcher's stated epistemological commitments. Under post-positivist and critical realist epistemologies, strategies emphasizing intersubjective validation member checking understood as verification, structural corroboration, and negative case analysis are most coherent. Under constructivist epistemologies, reflexivity, transparent documentation of researcher positionality, dialogic engagement with participants, and theoretical memoing are appropriate. Under post-structuralist and emancipatory epistemologies, Lather's (1986) concept of catalytic validity, alongside community feedback and transformative accountability, supplements or replaces traditional trustworthiness criteria because the research goals themselves are different in kind.

Dimension 3: Rigor Requirements

The third dimension specifies a minimum trustworthiness portfolio appropriate to each epistemological and methodological configuration. Some form of credibility strategy is required across all configurations; the appropriate strategy varies by epistemological stance. Dependability documentation through audit trails, reflexive journals, or analytical decision trails is specified as non-negotiable across all configurations, directly addressing the most consistent gap identified in the synthesis. Transferability is reconceptualized not as a property social science researchers can guarantee but as a judgment that readers including policymakers, practitioners, and community stakeholders make on the basis of contextual density provided. The researcher's responsibility is to offer thick description of social context, structural conditions, and participant

characteristics. Confirmability is addressed through transparent, accessible documentation of the analytic process, including iterations, discarded interpretations, and reflexive memos practices suited to open science repositories and institutional pre-registration platforms (Yardley et al., 2021).

Table 2

QMAF Configuration Matrix: Epistemological Stances, Aligned Traditions, and Rigor Strategies

Epistemological Stance	Aligned Traditions	Primary Rigor Mechanisms	Trustworthiness Focus	Key Indicators
Post-positivist / critical realism	Descriptive phenomenology, Yin case study, some grounded theory	Member checking (verification), structural corroboration, triangulation, negative case analysis	Credibility and dependability prioritized	Inter-coder reliability, rival explanations, audit trail
Constructivism	IPA, constructivist GT, narrative inquiry	Ongoing reflexivity, dialogic engagement, theoretical memoing, co-construction of meaning	Credibility and confirmability; dependability through reflexive journals	Researcher journal, analytical memos, member dialogue
Critical / post-structuralist / emancipatory	Critical ethnography, PAR, discourse analysis, feminist research	Catalytic validity, participant empowerment, transparency of power relations	Confirmability and transferability; outcomes-oriented and transformative accountability	Community feedback, structural analysis, transformative action
Pragmatism	Applied case study, ethnography, mixed	Multi-method triangulation, stakeholder review, fit-for-	All four criteria, weighted by context, audience, and	Practice relevance, stakeholder utility,

Epistemological Stance	Aligned Traditions	Primary Rigor Mechanisms	Trustworthiness Focus	Key Indicators
	qualitative	purpose evaluation	policy relevance	transferable findings

Note. GT = Grounded Theory; IPA = Interpretative Phenomenological Analysis; PAR = Participatory Action Research.

The QMAF is offered as a heuristic rather than a prescriptive algorithm. Social science researchers are encouraged to engage critically with its configurations, adapting them to the specific epistemological traditions, disciplinary conventions, and institutional contexts of their own work. Its primary value lies in making the epistemological assumptions underlying methodology selection explicit and open to scrutiny, a form of confirmability in itself.

Conclusion

This systematic review synthesized evidence from 94 qualitative social science studies to provide a comprehensive and critically grounded account of methodological rigor practices across the field. The picture that emerges is neither uniformly bleak nor unreservedly optimistic. Qualitative social science has made genuine progress in developing shared language for rigor, producing methodologically sophisticated exemplars, and generating a rich tradition of methodological scholarship across sociology, social work, education, and beyond. At the same time, the evidence shows that this language is applied inconsistently, that certain criteria are systematically underserved, that rigor practices often do not align with stated epistemological commitments, and that training in philosophical literacy has not kept pace with training in procedural technique.

The review's principal theoretical contribution, the QMAF, addresses these vulnerabilities by providing a structured pathway for aligning epistemological commitments, methodological traditions, and rigor strategies within the specific intellectual landscape of the social sciences. Unlike previous frameworks that have addressed traditions and criteria largely in isolation or within a single discipline, the QMAF integrates both within a coherent decision architecture applicable across social science domains.

Limitations

Several limitations of the present review should be acknowledged. First, the restriction to English-language publications introduces a geographic and cultural bias, likely excluding qualitative social science scholarship produced in

other linguistic traditions, a particularly significant constraint given the global scope of social science inquiry and the importance of non-Anglophone sociological and educational research traditions. The geographic distribution of included studies (North America, Europe, and Australasia accounting for 87% of the sample) means the findings may inadequately represent qualitative methodological practices in Global South social science research.

Second, the decision to exclude grey literature including dissertations, reports, and working papers may have introduced a positive publication bias, overrepresenting methodologically stronger work. This limitation is especially relevant in the social sciences, where policy reports and practitioner-oriented publications often contain important qualitative work that does not appear in indexed journals. Third, the operationalization of trustworthiness quality through textual content analysis captures what researchers report, not necessarily what they do. Fourth, the QMAF requires empirical testing across social science disciplines before its practical utility can be established. Finally, despite the multi-database approach, Booth's (2016) observation that qualitative social science research is highly dispersed across publication media means that some relevant work may have been missed even with the supplementary citation searching employed.

Future Directions

Future research directions include longitudinal analyses of rigor reporting trends in social science qualitative journals following the gradual adoption of ENTREQ and related reporting standards; experimental or qualitative evaluation of the QMAF's utility across different social science disciplines; cross-linguistic reviews incorporating qualitative scholarship beyond the English-language literature; and domain-specific meta-syntheses within political science qualitative methods, critical geography, or community-based participatory research traditions. Methodological advances in searching for qualitative social science evidence including theory-driven sampling approaches and improved database-specific search filters as documented by Booth (2016) should inform future reviews in this area. As qualitative social science continues to develop as a rigorous and socially consequential research tradition, the systematic and honest examination of its own practices is not an optional scholarly luxury. It is a foundational commitment to the integrity of knowledge about the social world.

Consent for publication

Not applicable

Availability of data and materials

The data supporting this study's findings are available on request from the corresponding author, TN. The data are not publicly available due to restrictions, e.g., they contain information that could compromise the privacy of research participants.

Competing interests

The author declare that there is no conflict of interest regarding the publication of this article. The author has no financial, personal, or professional relationships that could be perceived as influencing the content or conclusions of this work.

Funding

Not applicable

Authors' contributions

Thseen Nazir is solely responsible for all aspects of this research, including the conceptualization and design of the study, data collection, analysis, and interpretation. The author also conducted the literature review, developed the research methodology, and wrote the entire manuscript. All revisions and final approvals were completed by the author, ensuring the integrity and accuracy of the research presented.

Acknowledgments

The author would like to express gratitude to Ibn Haldun University for their valuable support in making this study possible.

References

- Birt, L., Scott, S., Cavers, D., Campbell, C., & Walter, F. (2016). Member checking: A tool to enhance trustworthiness or merely a nod to validation? *Qualitative Health Research*, 26(13), 1802–1811. <https://doi.org/10.1177/1049732316654870>
- Booth, A. (2016). Searching for qualitative research for inclusion in systematic reviews: A structured methodological review. *Systematic Reviews*, 5, Article 74. <https://doi.org/10.1186/s13643-016-0249-x>
- Bryman, A. (2016). *Social research methods* (5th ed.). Oxford University Press.

Bradshaw, C., Atkinson, S., & Doody, O. (2017). Employing a qualitative description approach in health care research. *Global Qualitative Nursing Research*, 4, 1–8. <https://doi.org/10.1177/2333393617742282>

Cahill, C. (2007). Doing research with young people: Participatory research and the rituals of collective work.

Children's Geographies, 5(3), 297–312.
<https://doi.org/10.1080/14733280701445895>

CASP. (2018). CASP qualitative checklist [2018]. Critical Appraisal Skills Programme. <https://casp-uk.net/casp-tools-checklists/>

Charmaz, K. (2014). *Constructing grounded theory* (2nd ed.). SAGE.

Cope, D. G. (2014). Methods and meanings: Credibility and trustworthiness of qualitative research. *Oncology Nursing Forum*, 41(1), 89–91.
<https://doi.org/10.1188/14.ONF.89-91>

Creswell, J. W., & Creswell, J. D. (2023). *Research design: Qualitative, quantitative, and mixed methods approaches* (6th ed.). SAGE.

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE.

Cypress, B. (2022). Qualitative research methods: A phenomenological focus. *Dimensions of Critical Care Nursing*, 41(2), 101–108.
<https://doi.org/10.1097/DCC.0000000000000493>

Denzin, N. K., & Lincoln, Y. S. (Eds.). (2018). *The SAGE handbook of qualitative research* (5th ed.). SAGE.

Dixon-Woods, M., Shaw, R. L., Agarwal, S., & Smith, J. A. (2007). The problem of appraising qualitative research. *Quality and Safety in Health Care*, 13(3), 223–225. <https://doi.org/10.1136/qshc.2003.008714>

Evans, D., & Pearson, A. (2001). Systematic reviews of qualitative research. *Clinical Effectiveness in Nursing*, 5(3), 111–119.
<https://doi.org/10.1054/cein.2001.0219>

Finlay, L. (2021). Thematic analysis: The 'good', the 'bad' and the 'ugly'. *European Journal for Qualitative Research in Psychotherapy*, 11, 103–116.

Flemming, K., Booth, A., Garside, R., Tuncalp, O., & Noyes, J. (2019). Qualitative evidence synthesis for complex interventions and guideline development. *BMJ Global Health*, 4(Suppl. 1), e000882.
<https://doi.org/10.1136/bmjgh-2018-000882>

Flyvbjerg, B. (2011). Case study. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (4th ed., pp. 301–316). SAGE.

Foley, G., & Timonen, V. (2015). Using grounded theory method to capture and analyze health care experiences. *Health Services Research*, 50(4), 1195–1210. <https://doi.org/10.1111/1475-6773.12275>

Giorgi, A. (2009). *The descriptive phenomenological method in psychology: A modified Husserlian approach*. Duquesne University Press.

Grayson, L., & Gomersall, A. (2003). *A difficult business: Finding the evidence for social science reviews*. ESRC UK Centre for Evidence Based Policy and Practice. Queen Mary, University of London.

Hammersley, M. (2008). Troubles with triangulation. In M. M. Bergman (Ed.), *Advances in mixed methods research* (pp. 22–36). SAGE.

Hannes, K., & Lockwood, C. (2012). *Synthesizing qualitative research: Choosing the right approach*. Wiley-Blackwell.

Haven, T. L., & Van Grootel, D. L. (2019). Preregistering qualitative research. *Accountability in Research*, 26(3), 229–244. <https://doi.org/10.1080/08989621.2019.1580147>

Kidd, S. A., & Kral, M. J. (2005). Practicing participatory action research. *Journal of Counseling Psychology*, 52(2), 187–195. <https://doi.org/10.1037/0022-0167.52.2.187>

Korstjens, I., & Moser, A. (2018). Series: Practical guidance to qualitative research. Part 4: Trustworthiness and publishing. *European Journal of General Practice*, 24(1), 120–124. <https://doi.org/10.1080/13814788.2017.1375092>

Lather, P. (1986). Issues of validity in openly ideological research: Between a rock and a soft place. *Interchange*, 17(4), 63–84. <https://doi.org/10.1007/BF01807017>

Lather, P. (2006). Paradigm proliferation as a good thing to think with: Teaching research in education as a wild profusion. *International Journal of Qualitative Studies in Education*, 19(1), 35–57. <https://doi.org/10.1080/09518390500450144>

Lietz, C. A., & Zayas, L. E. (2010). Evaluating qualitative research for social work practitioners. *Advances in Social Work*, 11(2), 188–202. <https://doi.org/10.18060/589>

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE.

Lloyd Jones, M. (2004). Application of systematic review methods to qualitative research: Practical issues. *Journal of Advanced Nursing*, 48(3), 271–278. <https://doi.org/10.1111/j.1365-2648.2004.03196.x>

Long, H. A., French, D. P., & Brooks, J. M. (2020). Optimising the value of the Critical Appraisal Skills Programme (CASP) tool for quality appraisal in qualitative evidence synthesis. *Research Methods in Medicine and Health Sciences*, 1(1), 31–42. <https://doi.org/10.1177/2632084320947559>

Morse, J. M. (2020). The changing face of qualitative inquiry. *International Journal of Qualitative Methods*, 19, 1–11. <https://doi.org/10.1177/1609406920909938>

Morse, J. M., Barrett, M., Mayan, M., Olson, K., & Spiers, J. (2002). Verification strategies for establishing reliability and validity in qualitative research. *International Journal of Qualitative Methods*, 1(2), 13–22.

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, 16(1), 1–13. <https://doi.org/10.1177/1609406917733847>

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan: A web and mobile app for systematic reviews. *Systematic Reviews*, 5(1), Article 210. <https://doi.org/10.1186/s13643-016-0384-4>

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hrobjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, Article n71. <https://doi.org/10.1136/bmj.n71>

Papaioannou, D., Sutton, A., Carroll, C., Booth, A., & Wong, R. (2010). Literature searching for social science systematic reviews: Consideration of a range of search techniques. *Health Information & Libraries Journal*, 27(2), 114–122.

Rolfe, G. (2006). Validity, trustworthiness and rigour: Quality and the idea of qualitative research. *Journal of Advanced Nursing*, 53(3), 304–310. <https://doi.org/10.1111/j.1365-2648.2006.03727.x>

Saldana, J. (2021). *The coding manual for qualitative researchers* (4th ed.). SAGE.

Shaw, R. L., Booth, A., Sutton, A., Miller, T., Smith, J., Young, B., & Dixon-Woods, M. (2004). Finding qualitative research: An evaluation of search strategies. *BMC Medical Research Methodology*, 4, Article 5.

Smith, B., & McGannon, K. R. (2018). Developing rigor in qualitative research: Problems and opportunities within sport and exercise psychology.

International Review of Sport and Exercise Psychology, 11(1), 101–121.
<https://doi.org/10.1080/1750984X.2017.1317357>

Smith, J. A., Flowers, P., & Larkin, M. (2009). Interpretative phenomenological analysis: Theory, method and research. SAGE.

Thomas, G. (2021). How to do your case study (3rd ed.). SAGE.

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Medical Research Methodology, 8(1), Article 45. <https://doi.org/10.1186/1471-2288-8-45>

Thornberg, R., & Charmaz, K. (2014). Grounded theory and theoretical coding. In U. Flick (Ed.), The SAGE handbook of qualitative data analysis (pp. 153–169). SAGE.

Tobin, G. A., & Begley, C. M. (2004). Methodological rigour within a qualitative framework. Journal of Advanced Nursing, 48(4), 388–396. <https://doi.org/10.1111/j.1365-2648.2004.03207.x>

Tong, A., Flemming, K., McInnes, E., Oliver, S., & Craig, J. (2022). Enhancing transparency in reporting the synthesis of qualitative research: ENTREQ. BMC Medical Research Methodology, 12(1), Article 181. <https://doi.org/10.1186/1471-2288-12-181>

Tracy, S. J. (2020). Qualitative research methods: Collecting evidence, crafting analysis, communicating impact (2nd ed.). Wiley-Blackwell.

Trainor, A. A., & Graue, E. (2013). Reviewing qualitative research in the social sciences. Routledge.

Yardley, L., Watts, S., Pearson, L., Muller, I., & Morrison, L. (2021). Ensuring that qualitative research is undertaken with appropriate standards of rigour. Psychology and Health, 36(10), 1173–1190. <https://doi.org/10.1080/08870446.2021.1977900>

Yin, R. K. (2018). Case study research and applications: Design and methods (6th ed.). SAGE.