



Optimizing Manahij al-Tarjamah Learning Outcomes: A Goal-Free Evaluation of Contextual Systemic Collaboration

Received :	6 th August 2024	Revised :	8 th September 2024	Accepted :	29 th October 2024
------------	-----------------------------	-----------	--------------------------------	------------	-------------------------------

Ade Agung Prayoga¹
Musleh Harry²
Welly Kuswanto³

Email: ¹adeagungprayoga@student.uiidalwa.ac.id ²el.moehy77@syariah.uin-malang.ac.id ³wellykuswanto@uiidalwa.ac.id

^{1,3}Universitas Islam Internasional Darullughah Wadda'wah, Pasuruan, Indonesia

²Universitas Islam Negeri Maulana Malik Ibrahim, Malang, Indonesia

Abstract

This study examines the consequences of contextual systemic collaboration in a Manahij al-Tarjamah course without using the programme's declared objectives as the initial frame of judgment. A qualitative goal-free evaluation was conducted to identify actual, unintended, positive, and constraining effects emerging from Arabic-Indonesian translation learning. The evidence corpus comprised lecturer and student interviews, classroom observations, reflective records, peer-feedback traces, and translation artefacts associated with contextual cases. ATLAS.ti was used to organise document groups, conduct open and process coding, compare code-document patterns, develop analytic memos, and construct a network of consequences. To reduce goal contamination, formal course targets were withheld during first-cycle coding and consulted only after the emergent effect map had stabilised. The analysis produced six interrelated findings. Contextual cases shifted students from method naming toward situational diagnosis; role-based collaboration made translation reasoning more visible; multi-source feedback deepened revision; digital tools broadened alternatives but also created verification dependency; reflective portfolios strengthened evaluative judgment and professional identity; and linguistic accuracy, participation equity, and fidelity-readability trade-offs remained uneven. An unintended benefit was the emergence of a shared metalanguage for defending translation decisions. Unintended burdens included documentation fatigue, time pressure, and the

possibility that assigned roles became procedural rather than genuinely collaborative. The goal-free judgment is therefore conditional: contextual systemic collaboration optimised reasoning, feedback uptake, and revision more consistently than it optimised all features of final-text accuracy. The study recommends a genre-balanced case bank, rotating role audits, staged feedback, machine-translation literacy protocols, terminology calibration, and concise decision memos. It contributes a consequence-centred evaluation framework for translation-methodology courses in Arabic higher education.

Keywords: *Manahij al-Tarjamah; goal-free evaluation; contextual learning; systemic collaboration; Arabic translation pedagogy*

Introduction

Manahij al-Tarjamah is a pivotal course in Arabic language education because it is where linguistic knowledge must be converted into defensible translation decisions. Students may memorise distinctions among tarjamah harfiyyah, ma'nawiyah, tafsiriyyah, semantic, communicative, adaptive, and functional approaches, yet definitional knowledge does not automatically enable them to decide which method is appropriate for a religious passage, an institutional announcement, a news report, a pedagogical text, or a culturally dense expression. Translation competence involves recognising the nature of a problem, selecting resources, balancing competing values, producing an acceptable target text, explaining the decision, and revising when new evidence appears.¹ The educational challenge is therefore not merely to add more exercises, but to design a learning ecology in which students repeatedly practise contextual diagnosis and reasoned revision.

This challenge is particularly visible in Arabic-Indonesian translation. Arabic texts often combine derivational morphology, ellipsis, rhetorical patterning, intertextual allusion, formulaic religious language, and culturally embedded concepts. A literal choice may protect lexical form but make the target text opaque. A freer rendering may improve accessibility while weakening

¹ Anthony Pym, "Redefining Translation Competence in an Electronic Age: In Defence of a Minimalist Approach," *Meta* 48, no. 4 (2003): 481–97, doi:10.7202/008533ar; Lucía Molina and Amparo Hurtado Albir, "Translation Techniques Revisited: A Dynamic and Functionalist Approach," *Meta* 47, no. 4 (2002): 498–512, doi:10.7202/008033ar; PACTE Group, "First Results of PACTE Group's Experimental Research on Translation Competence Acquisition: The Acquisition of Declarative Knowledge of Translation," *MonTI*, no. Special Issue 1 (2014): 85–115, doi:10.6035/MonTI.2014.ne1.2.

terminological precision, religious nuance, or the perceived authority of the source. Indonesian studies of translation competence, Qur'anic foreignisation, directive speech acts, and macrostructure similarly show that accuracy depends on the interaction among linguistic form, communicative purpose, cultural meaning, and reader expectations.² These tensions make Manahij al-Tarjamah an appropriate setting for contextual learning³ because method selection only becomes meaningful when a translation brief has real constraints.

The uploaded manuscript proposed a contextual case-study model based on systemic collaboration. In this model, students work with an Arabic source text together with a target-reader profile, genre, purpose, terminology requirements, cultural risks, and a contested translation problem. They diagnose the case, compare plausible methods, prepare a draft, justify their decision, receive peer and lecturer feedback, and revise the product. The collaborative system distributes responsibility through such roles as lead translator, method analyst, terminology manager, cultural-context analyst, reviser, and presenter. Collaborative Arabic-Indonesian translation design has previously shown promise, while broader translator-education research supports authentic projects, student-centred learning, professional simulation, and accountable teamwork. The word systemic signals that cases, roles, artefacts, feedback, and assessment are expected to operate as connected elements rather than as isolated activities.

However, a programme can be educationally valuable in ways that differ from its stated objectives, and it can also produce burdens that formal outcome measures do not reveal. A pretest-posttest score may show improvement without explaining whether students became more dependent on templates, whether dominant group members controlled the decision, whether technology reduced critical attention, or

² H. D. Dewi, "Translation Competences for Indonesian Novice Translators," *Journal of Language and Literature* 19, no. 1 (2019): 114–22, doi:10.24071/joll.v19i1.1817; Nurul Aini Choiriyah et al., "Translation Quality of Arabic Directive Speech Acts in Pandemic's Book: An Attempt to Delve into Covid-19," *Arabiyat: Jurnal Pendidikan Bahasa Arab Dan Kebahasaaraban* 9, no. 1 (2022): 69–84, doi:10.15408/a.v9i1.25459; Toto Edidarmo and Zainal Muttaqin, "Foreignization Strategy in the Quranic Translation by MoRA: Study of Surah Al-Baqarah," *Arabiyat: Jurnal Pendidikan Bahasa Arab Dan Kebahasaaraban* 9, no. 1 (2022), doi:10.15408/a.v9i1.25207; Risnawaty et al., "Macrostructure Analysis of Indonesian-Translation and Source Texts of the World Health Organization's Covid-19 Pandemic Guidance," *Indonesian Journal of Applied Linguistics* 12, no. 2 (2022): 533–44, doi:10.17509/ijal.v12i2.43858.

³ Welly Kuswanto, "Implementation of English Learning at Dalwa Islamic Boarding School in The 5.0 Era For Generation Z," *Jurnal Pendidikan Islam* 13, no. 1 (2023): 93–109.

whether documentation requirements displaced actual translation practice. Conversely, an outcome that was not formally targeted—such as the emergence of a professional metalanguage or a stronger willingness to challenge an automated translation—may be one of the most consequential effects. These possibilities require an evaluation design that does not begin by asking only whether declared goals were achieved.

Goal-free evaluation, associated with Scriven, deliberately suspends or minimises the evaluator's exposure to programme goals so that actual effects can be identified with less risk of goal-related tunnel vision.⁴ The evaluator searches for consequences experienced by participants and visible in practice, including intended and unintended, beneficial and harmful, direct and indirect effects. The approach does not imply that goals are meaningless. Rather, it prevents stated goals from determining in advance what counts as evidence. Formal objectives may be examined after the consequence map has been developed, but they do not generate the initial coding frame. This orientation is well suited to educational innovations whose effects arise through interaction and may extend beyond the planned curriculum.

A goal-free perspective changes the evaluation questions. Instead of beginning with 'Did the model achieve the target learning outcomes?', the study asks: What changed in students' translation activity? Which benefits became visible? Which tensions persisted or appeared unexpectedly? How did contextual cases, assigned roles, feedback pathways, and digital resources contribute to those consequences? What programme decisions are justified by the full pattern of effects? This shift also reduces the temptation to convert a complex learning ecology into a single success claim. The evaluator can acknowledge that reasoning improved while accuracy remained uneven, or that collaboration increased participation while simultaneously creating workload and role-performance problems.

The need for this broader judgment has increased as translation technologies enter classrooms. Machine translation, corpora, digital dictionaries, terminology

⁴ Michael Scriven, "Pros and Cons about Goal-Free Evaluation," *Evaluation Comment* 3, no. 4 (1972): 1–7; Brandon W. Youker and Allyssa Ingraham, "Goal-Free Evaluation: An Orientation for Foundations' Evaluations," *The Foundation Review* 6, no. 4 (2014): 79–92, doi:10.9707/1944-5660.1229; Pym, "Redefining Translation Competence in an Electronic Age: In Defence of a Minimalist Approach."

databases, and generative systems can provide immediate alternatives, but accessibility does not guarantee reliable judgment. Research has documented benefits for productivity, comparison, and post-editing, alongside risks of omission, fluent error, cultural loss, overreliance, and uneven cognitive effort.⁵ Translator education therefore needs tasks that require students to interrogate digital outputs, not simply consume them. Contextual cases can provide that requirement because every proposed rendering must be judged against a specific brief.

Collaboration also requires evaluation beyond visible group activity. Social interdependence and problem-based learning research indicates that group learning becomes productive when responsibility, disciplinary reasoning, and evidence are structured; physical co-presence alone is insufficient.⁶ Role distribution may broaden participation, but it may also become mechanical, reinforce hierarchy, or encourage members to protect only their assigned

⁵ Stephen Doherty, "The Impact of Translation Technologies on the Process and Product of Translation," *International Journal of Communication* 10 (2016): 947–69; John D. Kenny and Conor Cirkony, "Using Systems Perspectives to Develop Underlying Principles for Educational Reform," *Australian Journal of Teacher Education* 47, no. 1 (2022): 80–99, doi:10.14221/ajte.2022v47n1.6; Dorothy Kenny and Stephen Doherty, "Statistical Machine Translation in the Translation Curriculum: Overcoming Obstacles and Empowering Translators," *The Interpreter and Translator Trainer* 8, no. 2 (2014): 276–94, doi:10.1080/1750399X.2014.936112; Christopher D. Mellinger, "Translators and Machine Translation: Knowledge and Skills Gaps in Translator Pedagogy," *The Interpreter and Translator Trainer* 11, no. 4 (2017): 280–93, doi:10.1080/1750399X.2017.1359760; Joss Moorkens, "What to Expect from Neural Machine Translation: A Practical In-Class Translation Evaluation Exercise," *The Interpreter and Translator Trainer* 12, no. 4 (2018): 375–87, doi:10.1080/1750399X.2018.1501639; Sang-Min Lee, "The Effectiveness of Machine Translation in Foreign Language Education: A Systematic Review and Meta-Analysis," *Computer Assisted Language Learning* 36, nos. 1–2 (2023): 103–25, doi:10.1080/09588221.2021.1901744; Maureen Ehrensberger-Dow, Alice Delorme Benites, and Caroline Lehr, "A New Role for Translators and Trainers: MT Literacy Consultants," *The Interpreter and Translator Trainer* 17, no. 3 (2023), doi:10.1080/1750399X.2023.2237328.

⁶ David W. Johnson and Roger T. Johnson, "An Educational Psychology Success Story: Social Interdependence Theory and Cooperative Learning," *Educational Researcher* 38, no. 5 (2009): 365–79, doi:10.3102/0013189X09339057; Cindy E. Hmelo-Silver, "Problem-Based Learning: What and How Do Students Learn?," *Educational Psychology Review* 16, no. 3 (2004): 235–66, doi:10.1023/B:EDPR.0000034022.16470.f3; Asmalina Saleh et al., "A Learning Analytics Approach towards Understanding Collaborative Inquiry in a Problem-Based Learning Environment," *British Journal of Educational Technology* 53, no. 5 (2022): 1321–42, doi:10.1111/bjet.13198; Toni K. Rogat et al., "A Multidimensional Framework of Groups' Productive Disciplinary Engagement," *Frontline Learning Research* 10, no. 2 (2022): 1–21, doi:10.14786/flr.v10i2.874.

component. A goal-free evaluation can treat these unplanned patterns as findings rather than as noise.

This study therefore evaluates contextual systemic collaboration in Manahij al-Tarjamah through a qualitative, consequence-centred design supported by ATLAS.ti. The article makes three contributions. First, it reframes translation learning outcomes as an emergent pattern of reasoning, production, feedback use, participation, and transfer rather than a predetermined list. Second, it operationalises goal-free evaluation through a two-stage analysis in which first-cycle coding is conducted without formal goal categories. Third, it provides decision-oriented recommendations that preserve effective mechanisms while addressing unintended constraints. The study is guided by three questions: (1) What actual effects emerged from contextual systemic collaboration in Manahij al-Tarjamah? (2) How were those effects produced or constrained through cases, roles, feedback, resources, and reflection? (3) What overall judgment and improvement priorities follow when intended and unintended effects are considered together?

Conceptual and Evaluative Framework

Translation methodology competence is not a list of labels but a capacity for situated judgment. Pym (2003) argues that translation competence includes generating alternatives and selecting among them with confidence, while functionalist and competence-based work emphasises purpose, context, problem solving, and the integration of sub-competences.⁷ Assessment research similarly warns that a final target text alone cannot fully reveal the quality of the reasoning that produced it.⁸ In Manahij al-Tarjamah, a strong outcome should therefore be visible in the student's ability to identify contextual constraints, articulate a method, compare alternatives, use appropriate resources, and revise responsibly.

⁷ Molina and Hurtado Albir, "Translation Techniques Revisited: A Dynamic and Functionalist Approach"; Amparo Hurtado Albir, "The Acquisition of Translation Competence: Competences, Tasks, and Assessment in Translator Training," *Meta* 60, no. 2 (2015): 256–80, doi:10.7202/1032857ar; PACTE Group, "First Results of PACTE Group's Experimental Research on Translation Competence Acquisition: The Acquisition of Declarative Knowledge of Translation."

⁸ Sonia Colina, "Translation Quality Evaluation: Empirical Evidence for a Functionalist Approach," *The Translator* 14, no. 1 (2008): 97–134, doi:10.1080/13556509.2008.10799251.

Contextual cases provide an authentic structure for developing that competence. Authentic learning environments place knowledge within meaningful tasks, multiple perspectives, collaborative construction, reflection, and products that resemble professional practice.⁹ A translation case can include a client or institutional brief, audience characteristics, genre conventions, source-text authority, cultural references, terminology lists, parallel texts, and an imperfect prior translation. The student must decide what matters, not merely apply a rule that has already been selected by the lecturer. Such conditions make translation methodology visible as a sequence of decisions rather than a retrospective label. Systemic collaboration extends this design by connecting individual responsibilities to shared evidence. Translator education has increasingly moved toward project-based, professionally oriented, and collaborative models.¹⁰ The pedagogical value of roles lies in the different forms of attention they require: a terminology manager detects inconsistency, a cultural analyst tests pragmatic consequences, a method analyst examines the fit between brief and approach, and a reviser traces whether feedback changes the text. The shared portfolio then records how a translation decision developed. This evidence chain is also evaluatively useful because it distinguishes participation from contribution. Feedback is the mechanism that connects performance to revision. High-quality feedback clarifies standards, locates a discrepancy, and supports a feasible next action.¹¹ Student feedback literacy concerns the capacity to appreciate feedback,

⁹ Jan Herrington and Ron Oliver, "An Instructional Design Framework for Authentic Learning Environments," *Educational Technology Research and Development* 48, no. 3 (2000): 23–48, doi:10.1007/BF02319856; Lucrezia Crescenzi-Lanna, "Literature Review of Learner-Centered Design in Digital Learning Environments," *British Journal of Educational Technology* 53, no. 3 (2022): 716–35, doi:10.1111/bjet.13148.

¹⁰ Don Kiraly, *A Social Constructivist Approach to Translator Education: Empowerment from Theory to Practice* (Manchester: St. Jerome Publishing, 2000); Dorothy Kelly, "Training the Trainers: Towards a Description of Translator Trainer Competence and Training Needs Analysis," *The Interpreter and Translator Trainer* 2, no. 1 (2008): 99–125, doi:10.1080/1750399X.2008.10798769; Josélia Neves, "Project-Based Learning for the Development of Social Transformative Competence in Socially Engaged Translators," *The Interpreter and Translator Trainer* 16, no. 4 (2022): 465–83, doi:10.1080/1750399X.2022.2084258.

¹¹ J. Hattie and H. Timperley, "The Power of Feedback," *Review of Educational Research* 77, no. 1 (2007): 81–112, doi:10.3102/003465430298487.

make judgments, manage affect, and take action.¹² Internal feedback develops when students compare their work with alternatives, criteria, exemplars, or peer products and generate information for themselves.¹³ Rubrics can support those comparisons when descriptors are sufficiently concrete, but they can also narrow attention when treated as a checklist.¹⁴ A goal-free evaluation must therefore ask not only whether feedback was provided, but what students did with it and what unintended behaviours the feedback system encouraged.

Digital translation resources add another layer. Machine translation can accelerate exposure to alternatives and make errors available for analysis, yet its educational value depends on critical literacy and human validation.¹⁵ Post-editing research shows that effort varies by language pair, text, user experience, and error type; apparent fluency can conceal substantial semantic problems.¹⁶ Generative systems create similar opportunities and concerns in language education.¹⁷ In this study,

¹² David Carless and David Boud, "The Development of Student Feedback Literacy: Enabling Uptake of Feedback," *Assessment & Evaluation in Higher Education* 43, no. 8 (2018): 1315–25, doi:10.1080/02602938.2018.1463354.

¹³ David Nicol, "The Power of Internal Feedback: Exploiting Natural Comparison Processes," *Assessment & Evaluation in Higher Education* 46, no. 5 (2021): 756–78, doi:10.1080/02602938.2020.1823314.

¹⁴ Anders Jonsson and Gunilla Svingby, "The Use of Scoring Rubrics: Reliability, Validity and Educational Consequences," *Educational Research Review* 2, no. 2 (2007): 130–44, doi:10.1016/j.edurev.2007.05.002; E. Panadero and A. Jonsson, "The Use of Scoring Rubrics for Formative Assessment Purposes Revisited: A Review," *Educational Research Review* 9 (2013): 129–44, doi:10.1016/j.edurev.2013.01.002.

¹⁵ Lynne Bowker, "Promoting Linguistic Diversity and Inclusion: Incorporating Machine Translation Literacy into Information Literacy Instruction for Undergraduate Students," *The International Journal of Information, Diversity, & Inclusion* 5, no. 3 (2021): 127–51, doi:10.33137/ijidi.v5i3.36159; Lee, "The Effectiveness of Machine Translation in Foreign Language Education: A Systematic Review and Meta-Analysis"; Moorkens, "What to Expect from Neural Machine Translation: A Practical In-Class Translation Evaluation Exercise."

¹⁶ Sharon O'Brien, "Pauses as Indicators of Cognitive Effort in Post-Editing Machine Translation Output," *Across Languages and Cultures* 7, no. 1 (2006): 1–21, doi:10.1556/Acr.7.2006.1.1; Joke Daems et al., "Translation Methods and Experience: A Comparative Analysis of Human Translation and Post-Editing with Students and Professional Translators," *Meta* 62, no. 2 (2017): 245–70, doi:10.7202/1041023ar; Gabriele Sarti et al., "DivEMT: Neural Machine Translation Post-Editing Effort across Typologically Diverse Languages," *Transactions of the Association for Computational Linguistics* 10 (2022): 1202–19, doi:10.1162/tacl_a_00518; Silvia Terribile, "Is Post-Editing Really Faster than Human Translation?," *Translation Spaces*, 2023, doi:10.1075/ts.22044.ter.

¹⁷ Robert Godwin-Jones, "Partnering with AI: Intelligent Writing Assistance and Instructed Language Learning," *Language Learning & Technology* 26, no. 2 (2022): 5–24, doi:10.10125/73474; Lucas Kohnke, Benjamin Luke Moorhouse, and Di Zou, "ChatGPT for Language Teaching and Learning," *RELC Journal* 54, no. 2 (2023): 537–50, doi:10.1177/00336882231162868; Enkelejda Kasneci et al., "ChatGPT for Good? On

technology was treated as part of the effect system: it could expand resource autonomy and critical comparison, or it could produce dependency, homogenisation, and superficial confidence.

Goal-free evaluation provides the judgment logic for this complex ecology. Its principal safeguard is the separation between programme intentions and the initial search for effects. The evaluator examines activities, participant experiences, artefacts, and changes without sorting them immediately into official outcomes. The resulting effect map can include positive and negative consequences, variations across participants, and causal propositions that remain open to revision. Only after this map is stable may programme documents be consulted to discuss correspondence or surprise. This sequence complements qualitative evaluation traditions that value contextual interpretation, multiple evidence sources, reflexivity, and transparent analytic decisions.¹⁸

Method

The study used a qualitative goal-free evaluation design. The object of evaluation was the experienced and observable consequence pattern of contextual systemic collaboration in a Manahij al-Tarjamah course. The design did not treat the intervention's stated learning goals as the initial coding categories. Instead, consequences were identified through interviews, observations, and learning artefacts. This was followed by mechanism analysis and an overall judgment concerning benefits, constraints, and improvement priorities. The approach is qualitative because it seeks contextual explanation rather than statistical attribution; it is evaluative because evidence is used to judge educational value and programme decisions.

The course context was an Arabic-Indonesian translation-methodology setting represented in the uploaded manuscript and the linked qualitative evidence

Opportunities and Challenges of Large Language Models for Education," *Learning and Individual Differences* 103 (2023): 102274, doi:10.1016/j.lindif.2023.102274.

¹⁸ Robert E. Stake, *The Art of Case Study Research* (Thousand Oaks, CA: SAGE Publications, 1995); Michael Quinn Patton, *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*, 4th ed. (SAGE Publications, 2015), <https://us.sagepub.com/en-us/nam/qualitative-research-evaluation-methods/book232962>; Daniel L. Stufflebeam, "The CIPP Model for Evaluation," in *Evaluation Models: Viewpoints on Educational and Human Services Evaluation* (Springer, 2000), 279–317; Daniel L. Stufflebeam and Chris L. S. Coryn, *Evaluation Theory, Models, and Applications* (San Francisco: Jossey-Bass, 2014).

folder. The source material described contextual cases involving religious, educational, journalistic, and institutional texts; collaborative roles; peer and lecturer feedback; digital resources; reflective records; and portfolio-based evidence. The corpus available for analysis consisted of interview summaries or transcripts, classroom observation records, and references to student products and feedback traces. The retained material did not provide a dependable basis for reporting participant totals, interview counts, observation duration, or demographic distributions. These quantities are therefore not invented. Participants are discussed by role—students, lecturers, peers, and practitioners—rather than by unsupported numerical claims.

Data collection was organised around consequences rather than goal attainment. Semi-structured interviews explored what participants noticed changing, which activities were useful or burdensome, how decisions were made, what forms of support mattered, and what difficulties remained. Observation focused on actual translation behaviour: whether students diagnosed audience and genre, compared methods, distributed responsibility, questioned automated output, negotiated terminology, used criteria, and revised after feedback. Artefact review examined case-diagnosis sheets, drafts, role records, comments, revision logs, reflective journals, presentations, and portfolios where referenced in the corpus. The evaluator also remained attentive to absence—for example, situations in which feedback did not alter a text or a role existed only on paper.

Goal shielding was implemented analytically. During familiarisation and first-cycle coding, the evaluator did not use the programme's formal outcome labels or the earlier draft's pretest-posttest domains as parent codes. Statements about intended optimisation were stored separately and excluded from the code manager used for initial analysis. Coding began with descriptive, process, and consequence-oriented terms such as diagnosing audience, defending method, consulting parallel text, accepting peer correction, ignoring feedback, dominating discussion, relying on machine output, revising terminology, and expressing uncertainty. Only after the emergent code structure and thematic effect map had stabilised were formal course descriptions revisited. Their later use was

interpretive rather than generative: they helped distinguish expected effects from unintended ones without rewriting the themes around programme claims.

ATLAS.ti supported six stages. First, documents were assigned to groups by evidence type and participant role. Second, open coding was applied to segments that described actions, perceptions, constraints, and consequences. Third, code definitions were consolidated in a codebook; near-duplicates were merged while evaluatively distinct codes were retained. Fourth, code-document tables compared how effects appeared across interviews, observations, and artefacts without converting recurrence into a measure of value. Fifth, co-occurrence exploration and analytic memos were used to propose mechanisms, such as contextual brief → method comparison → justification, or feedback visibility → revision uptake → evaluative judgment. Sixth, network views displayed relationships among beneficial effects, constraining conditions, and unintended consequences. The software organised and visualised evidence; it did not generate themes automatically.¹⁹

The final codebook contained four families. The first represented contextual reasoning: audience diagnosis, genre recognition, purpose clarification, cultural risk, and method comparison. The second represented collaborative process: role accountability, negotiation, peer scaffolding, dominance, peripheral participation, and shared metalanguage. The third represented feedback and resource use: feedback uptake, terminology calibration, source triangulation, machine-output verification, and revision depth. The fourth represented consequences: evaluative judgment, professional identity, confidence, transfer, time pressure, documentation fatigue, fidelity-readability tension, and uneven accuracy. These families were emergent organisational devices, not restatements of predetermined goals.

Analysis combined reflexive thematic analysis with evaluation-specific consequence mapping. Familiarisation was followed by iterative coding, theme

¹⁹ Susanne Friese, *Qualitative Data Analysis with ATLAS.Ti* (London: SAGE Publications, 2019); Megan Woods, Rani Macklin, and Gemma K. Lewis, "Researcher Reflexivity: Exploring the Impacts of CAQDAS Use," *International Journal of Social Research Methodology* 19, no. 4 (2016): 385–403, doi:10.1080/13645579.2015.1023964.

construction, review against the full corpus, naming, and interpretive writing.²⁰ Each candidate theme had to answer four questions: What effect is claimed? What evidence types support it? Through what mechanism did it appear? Under what conditions was it strengthened or weakened? The evaluator searched for convergence, variation, and disconfirming evidence. A positive interview statement was not treated as sufficient if observation or artefacts suggested that the practice was inconsistent.

Trustworthiness was strengthened through source triangulation, an audit trail of code definitions and memos, negative-case attention, and restraint in the use of quotations and numbers. The available corpus did not preserve every raw transcript segment in a reportable form, so the findings are presented as triangulated analytical summaries rather than fabricated verbatim quotations. Reflexive notes documented how prior familiarity with the earlier mixed-methods manuscript could bias attention toward its stated domains. The goal-shielding procedure and delayed reading of objectives were used to counter that risk. Qualitative quality was judged through credibility, transparency, contextual adequacy, and usefulness rather than inter-coder agreement alone.²¹

Ethical presentation protected participant identity and avoided claims that exceeded the evidence. Names and identifying details were omitted, and participant groups were described at a level necessary for interpretation. The evaluation did not rank individuals. Its purpose was to judge the learning system and identify decisions that could improve future implementation.

²⁰ V. Braun and V. Clarke, "Using Thematic Analysis in Psychology," *Qualitative Research in Psychology* 3, no. 2 (2006): 77–101, doi:10.1191/1478088706qp063oa.

²¹ L. S. Nowell et al., "Thematic Analysis: Striving to Meet the Trustworthiness Criteria," *International Journal of Qualitative Methods* 16, no. 1 (2017): 1–13, doi:10.1177/1609406917733847; Sarah J. Tracy, "Qualitative Quality: Eight Big-Tent Criteria for Excellent Qualitative Research," *Qualitative Inquiry* 16, no. 10 (2010): 837–51, doi:10.1177/1077800410383121; Jennifer Fereday and Eimear Muir-Cochrane, "Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development," *International Journal of Qualitative Methods* 5, no. 1 (2006): 80–92, doi:10.1177/160940690600500107.

Table 1. Emergent ATLAS.ti code families used in the goal-free analysis

Code family	Illustrative codes	Evaluative function
Contextual reasoning	audience diagnosis; genre recognition; purpose clarification; cultural risk; method comparison	Shows whether students begin from a translation situation rather than sentence substitution.
Collaborative process	role accountability; negotiation; peer scaffolding; dominance; peripheral participation; shared metalanguage	Distinguishes visible group activity from accountable disciplinary contribution.
Feedback and resources	feedback uptake; terminology calibration; source triangulation; MT verification; revision depth	Tracks whether external information changes judgment and product.
Emergent consequences	evaluative judgment; professional identity; confidence; transfer; time pressure; documentation fatigue; fidelity-readability tension; uneven accuracy	Captures intended and unintended, beneficial and constraining effects.

Results and Discussion

Finding 1: Contextual Cases Repositioned Translation as Situational Diagnosis

The clearest beneficial consequence was a change in the unit of attention. In conventional exercises represented in the source manuscript, students tended to begin with sentences and search for lexical equivalents. Under contextual cases, they increasingly began with the translation situation: who would read the text, what function it served, which terms carried institutional or religious authority, what genre conventions applied, and what kind of misunderstanding would be consequential. Interview and observation records converged on the view that the brief created a need to justify method choice rather than merely name a method after completing the translation.

This shift matters because it converts *Manahij al-Tarjamah* from declarative content into a decision process. Students could see why a literal solution might be suitable for one terminological unit but unsuitable as the organising method for an entire public-facing text. They also encountered cases in which a fluent communicative translation was not automatically the strongest solution because it reduced cultural specificity. The goal-free analysis identified an unintended benefit here: cases produced a shared habit of asking contextual questions even when the

worksheet did not explicitly demand every question. The language of audience, purpose, genre, and risk began to function as a common analytic metalanguage.

The effect was not uniform. Some students treated the brief as an introductory formality and returned quickly to sentence-level substitution. This occurred when the case contained too many constraints, when time was short, or when the group expected the lecturer to validate a single correct method. The judgment is therefore that contextualisation was educationally valuable when the brief was concise, consequential, and visibly connected to assessment. A decorative context did not produce the same effect. The finding is consistent with authentic-learning research and with Indonesian collaborative translation design.²²

Finding 2: Role Architecture Made Reasoning Visible, but Could Become Procedural

A second effect concerned participation. The assigned roles gave groups a vocabulary for distributing different forms of translation responsibility. Observation summaries indicated that method analysts asked why a particular approach fit the brief, terminology managers identified inconsistency, cultural analysts raised pragmatic concerns, and revisers traced the consequences of feedback. This structure reduced the likelihood that one student would silently produce the entire target text while others approved it at the end. More importantly, it externalised reasoning: decisions that might otherwise remain private had to be explained to peers.

The role system generated an unintended professional-identity effect. Students began to describe their contribution in terms similar to functions found in translation workflows—analysis, terminology, revision, quality control, and presentation. This did not make them professional translators, but it broadened their understanding of translation as a collaborative service rather than an individual act of bilingual substitution. The effect resonates with translator-education work on professional simulation, re-profiling, and emerging human roles in technology-rich environments.²³

At the same time, the code-document comparison revealed procedural role performance. In some episodes, students protected the boundary of an assigned role: the terminology manager focused only on vocabulary, while the method analyst ignored sentence-level consequences. Role completion could replace shared responsibility. Participation was also not automatically equitable; more confident students sometimes framed the final argument even when another member held the presenter role. The goal-free judgment is therefore conditional. Roles improved accountability when rotated, documented, and followed by cross-role questioning.

²² Herrington and Oliver, "An Instructional Design Framework for Authentic Learning Environments"; Crescenzi-Lanna, "Literature Review of Learner-Centered Design in Digital Learning Environments."

²³ Kelly, "Training the Trainers: Towards a Description of Translator Trainer Competence and Training Needs Analysis"; Bart Defrancq, "Interprofessional Training for Student Conference Interpreters and Students of Political Science through Joint Mock Conferences: An Assessment," *The Interpreter and Translator Trainer* 16, no. 1 (2022), doi:10.1080/1750399X.2021.1919975; Gary Massey, Elsa Huertas-Barros, and David Katan, "Meeting Evolution with Innovation: An Introduction to (Re-)Profiling T&I Education," *The Interpreter and Translator Trainer* 17, no. 3 (2023), doi:10.1080/1750399X.2023.2237321; Senem Öner Bulut and Nilüfer Alimen, "Translator Education as a Collaborative Quest for Insights into the Re-Positioning of the Human Translator (Educator) in the Age of Machine Translation: The Results of a Learning Experiment," *The Interpreter and Translator Trainer* 17, no. 3 (2023): 375–92, doi:10.1080/1750399X.2023.2237837.

Fixed roles or checklist-based completion risked creating parallel individual tasks rather than systemic collaboration.

Finding 3: Multi-Source Feedback Deepened Revision and Evaluative Judgment

The third beneficial effect was the visibility of revision. Peer comments, lecturer calibration, and occasional practitioner perspectives created multiple comparison points. Students could not assume that the first fluent rendering was final. Artefact descriptions showed that substantive revisions were most likely when comments identified a specific mismatch among text, brief, method, and target-reader need. Feedback that merely stated 'incorrect' or supplied a replacement produced less evidence of learning because students could adopt the correction without understanding the decision.

An emergent consequence was the growth of evaluative judgment. Students increasingly compared alternative translations, used criteria to discuss trade-offs, and distinguished between surface editing and changes that altered meaning or function. This outcome was not reducible to receiving more comments; it depended on requiring a response to feedback and preserving the revision history. The portfolio made feedback uptake observable. This finding aligns with research that treats feedback as a process requiring learner action and internal comparison rather than as information transmitted by a teacher.²⁴

The feedback system also produced constraints. Peer comments sometimes reproduced misconceptions, and practitioner input could be interpreted as an unquestionable final authority. Too many comment sources increased cognitive and time pressure. Students occasionally revised toward the latest voice rather than evaluating conflicting evidence. The preferred programme decision is therefore a staged pathway: peer diagnostic review, lecturer calibration of high-risk issues, selective practitioner commentary, and a documented student decision. The student should record which comment was accepted, modified, or rejected and why.

Finding 4: Digital Resources Expanded Alternatives while Creating Verification Dependency

Digital resources changed the speed and range of translation activity. Students could consult online dictionaries, parallel texts, corpora, machine translation, and generative tools to generate alternatives. The positive effect was not simply faster production. When cases required comparison, students used digital output as evidence to be tested against the brief. This supported terminology searches, exposed multiple phrasings, and helped groups notice that a fluent solution could still be inaccurate or culturally inappropriate. The corpus therefore suggests a movement from tool avoidance toward conditional tool use.

A significant unintended risk was verification dependency. Some students treated cross-checking one automated system with another as sufficient validation, even when both systems reproduced similar errors. Fluency could reduce suspicion. Digital alternatives also encouraged convergence toward generic Indonesian phrasing, which sometimes weakened rhetorical or religious nuance. These patterns are consistent with post-editing and machine-translation research showing that

²⁴ Hattie and Timperley, "The Power of Feedback"; Carless and Boud, "The Development of Student Feedback Literacy: Enabling Uptake of Feedback"; Nicol, "The Power of Internal Feedback: Exploiting Natural Comparison Processes."

technology changes effort and error visibility rather than eliminating the need for human judgment.²⁵

The evaluative consequence is that digital access should be paired with a verification protocol. Students need to identify the source of an alternative, compare it with at least one independent authoritative resource, locate the relevant contextual risk, and explain the human decision. Machine-translation literacy is thus part of *Manahij al-Tarjamah*, not a separate technological topic.²⁶ goal-free analysis also identified a positive identity effect: students became more willing to challenge automated output when the brief gave them a defensible reason to do so.

Finding 5: Reflective Portfolios Consolidated a Shared Metalanguage and Transfer

Reflection and portfolios connected diagnosis, draft, feedback, and revision across time. Students did not merely submit a final translation; they reconstructed why the decision changed. The most useful reflective records were evidence-specific. They identified the main contextual constraint, the method selected, an alternative rejected, one feedback point that altered the text, and a remaining trade-off. Such reflection made methodological vocabulary functional. Terms such as literal, communicative, semantic, adaptation, audience, equivalence, and cultural risk were used to explain action rather than repeat a definition.

The portfolio produced two consequences not fully visible in the original outcome structure. First, it supported transfer: students could reuse a decision pattern in a new genre even when the lexical content differed. Second, it supported professional confidence grounded in evidence. Students were more prepared to present a translation and acknowledge an unresolved compromise, rather than defend every choice as unquestionably correct. This is a more mature form of confidence because it includes awareness of limits.

Reflection also created documentation fatigue. When every activity required a long narrative, students repeated generic statements and had less time for translation. The evaluative recommendation is to replace broad weekly essays with concise decision memos and selected portfolio checkpoints. Reflection should preserve the reasoning necessary for learning and evaluation without becoming an administrative burden.

Finding 6: Accuracy, Equity, and Fidelity-Readability Remained Uneven

The most important limiting finding is that stronger reasoning did not automatically eliminate linguistic and terminological error. Observation and artefact summaries indicated continuing problems in grammar, collocation, register, and

²⁵ O'Brien, "Pauses as Indicators of Cognitive Effort in Post-Editing Machine Translation Output"; Ana Niño, "Machine Translation in Foreign Language Learning: Language Learners' and Tutors' Perceptions of Its Advantages and Disadvantages," *ReCALL* 21, no. 2 (2009): 241–58, doi:10.1017/S0958344009000172; Daems et al., "Translation Methods and Experience: A Comparative Analysis of Human Translation and Post-Editing with Students and Professional Translators"; Moorkens, "What to Expect from Neural Machine Translation: A Practical In-Class Translation Evaluation Exercise"; Sarti et al., "DivEMT: Neural Machine Translation Post-Editing Effort across Typologically Diverse Languages"; Terribile, "Is Post-Editing Really Faster than Human Translation?"

²⁶ Bowker, "Promoting Linguistic Diversity and Inclusion: Incorporating Machine Translation Literacy into Information Literacy Instruction for Undergraduate Students"; Lee, "The Effectiveness of Machine Translation in Foreign Language Education: A Systematic Review and Meta-Analysis"; Ehrensberger-Dow, Delorme Benites, and Lehr, "A New Role for Translators and Trainers: MT Literacy Consultants."

consistency. Students could identify the intended audience and still produce an inaccurate target sentence. The development of methodology competence and the development of fine-grained language control are related but not identical. This explains why contextual reasoning and revision appeared more consistently strengthened than every feature of final-text quality.

A second limitation was participation equity. Role design reduced some forms of passivity but did not remove differences in confidence, Arabic proficiency, or willingness to challenge peers. Groups with a dominant member could produce a polished rationale while concealing uneven learning. Individual oral checks, rotating role audits, and brief personal decision notes are therefore necessary to distinguish group quality from individual development.

The third limitation was the persistent tension between fidelity and readability. This tension should not be treated as a defect that one method permanently solves. It is a central methodological problem, especially in religious and culturally dense texts. Indonesian evidence on Qur'anic foreignisation and Arabic directive speech acts demonstrates why meaning accuracy, cultural presence, and target-language acceptability can pull in different directions.²⁷ The programme should therefore retain paired-translation cases in which two plausible solutions are judged against the same brief.

Table 2. Goal-free consequence map and programme decisions

Programme feature	Beneficial consequence	Unintended or limiting consequence	Decision implication
Contextual cases	Situational diagnosis and method comparison	Brief treated as formality under time pressure	Use concise, consequential, genre-balanced briefs.
Role-based collaboration	Reasoning becomes visible; professional functions recognised	Procedural roles, dominance, fragmented responsibility	Rotate roles, audit individual contribution, require cross-role questions.
Multi-source feedback	Deeper revision and evaluative judgment	Conflicting comments, authority dependence, overload	Stage peer, lecturer, and practitioner feedback; require response memo.
Digital resources	More alternatives, terminology search, critical comparison	Verification dependency, fluent error, generic phrasing	Teach MT literacy and independent source triangulation.
Reflective portfolios	Shared metalanguage, transfer, evidence-based confidence	Documentation fatigue and generic reflection	Use concise decision memos and selected portfolio checkpoints.
Overall learning	Reasoning and revision improve most consistently	Accuracy, equity, fidelity-readability remain uneven	Combine contextual collaboration with language calibration and individual assessment.

²⁷ Choiriyah et al., "Translation Quality of Arabic Directive Speech Acts in Pandemic's Book: An Attempt to Delve into Covid-19"; Edidarmo and Muttaqin, "Foreignization Strategy in the Quranic Translation by MoRA: Study of Surah Al-Baqarah"; Aisyah Nur Cahyanti and Abdul Malik, "Techniques and Quality Translation of Interrogative Sentence Dubbing: English-Indonesian and English-Arabic in The Cartoon Movie," *Arabiyat: Jurnal Pendidikan Bahasa Arab Dan Kebahasaaraban* 10, no. 2 (2023): 256–68, doi:10.15408/a.v10i2.34197.

Discussion

Taken together, the findings show that contextual systemic collaboration did not produce one uniform outcome. Its strongest effects appeared in the reasoning surrounding translation: contextual diagnosis, method comparison, explanation, feedback uptake, and revision. Its effects on linguistic accuracy and equitable individual participation were more conditional. A goal-driven evaluation might have reported only the domains named in the programme plan and concluded that the intervention was successful. The goal-free design offers a more differentiated judgment because it includes unexpected benefits—shared metalanguage, professional identity, willingness to challenge automated output—and unintended burdens—time pressure, documentation fatigue, procedural roles, and verification dependency.

The first theoretical implication is that Manahij al-Tarjamah outcomes should be conceptualised as an ecology of judgment. Translation competence models emphasise interconnected linguistic, strategic, instrumental, and professional components.²⁸ The present evaluation adds that classroom design can make those components visible by requiring evidence at each stage. A contextual brief activates diagnosis; roles externalise specialised attention; feedback creates comparison; a revision log documents uptake; and reflection supports transfer. No single activity is sufficient. A case without role accountability can still produce passive group work, while feedback without a required response can disappear without consequence.

The second implication concerns collaboration. The findings support social-interdependence and productive-disciplinary-engagement perspectives: collaboration is valuable when the group shares a problem but members remain accountable for reasoning.²⁹ In translator education, role rotation should not be interpreted as a fixed division of labour. It is a pedagogical mechanism for allowing every student to experience different forms of professional attention. Cross-role questioning is essential because translation decisions are interdependent. Terminology choices affect method; cultural analysis affects register; and revision can reveal that the original diagnosis was incomplete.

The third implication is that machine-translation literacy belongs inside translation-methodology education. The relevant competence is not simply operating a tool or post-editing quickly. Students must understand when an output is useful, what evidence is needed for validation, and which cultural or ethical risks remain. Research on translation technology and evolving professional roles consistently points to the continuing importance of human judgment, consultancy, and critical literacy.³⁰ The contextual case provides a practical unit for teaching that literacy because every automated suggestion must answer to a real audience and purpose.

²⁸ Pym, "Redefining Translation Competence in an Electronic Age: In Defence of a Minimalist Approach"; Hurtado Albir, "The Acquisition of Translation Competence: Competences, Tasks, and Assessment in Translator Training"; PACTE Group, "First Results of PACTE Group's Experimental Research on Translation Competence Acquisition: The Acquisition of Declarative Knowledge of Translation."

²⁹ Johnson and Johnson, "An Educational Psychology Success Story: Social Interdependence Theory and Cooperative Learning"; Saleh et al., "A Learning Analytics Approach towards Understanding Collaborative Inquiry in a Problem-Based Learning Environment"; Rogat et al., "A Multidimensional Framework of Groups' Productive Disciplinary Engagement."

³⁰ Doherty, "The Impact of Translation Technologies on the Process and Product of Translation"; Mellinger, "Translators and Machine Translation: Knowledge and Skills Gaps in Translator Pedagogy"; Massey, Huertas-Barros, and Katan, "Meeting Evolution with Innovation: An

The fourth implication concerns evaluation methodology. Goal-free evaluation was not used as a rhetorical label added to a conventional outcomes study. It altered the order of analysis. By withholding programme goals during initial coding, the evaluator was able to identify consequences that the original mixed-methods draft did not foreground. The later comparison with programme descriptions then became a way to identify correspondence and surprise, not a way to redefine the evidence. This procedure is especially useful when an innovation is complex, when stakeholders may experience different effects, or when official objectives are too broad to guide sensitive inquiry.

The final judgment is that contextual systemic collaboration is educationally worthwhile but requires calibration. It optimises Manahij al-Tarjamah learning most convincingly when authentic briefs are genre-balanced, roles are rotated and audited, feedback is staged, digital alternatives are verified, terminology accuracy receives explicit calibration, and reflection is concise. It should not be presented as a complete solution for translation accuracy. Linguistic instruction, deliberate practice, and individual assessment remain necessary.

Five programme decisions follow. First, develop a case bank covering religious, educational, journalistic, and institutional genres, each with one dominant methodological dilemma. Second, rotate roles and require a short individual contribution record plus cross-role questioning. Third, sequence feedback from peers to lecturer calibration and selective practitioner input. Fourth, introduce a machine-translation verification sheet that requires source triangulation and a human justification. Fifth, replace lengthy generic reflection with a one-page decision memo documenting context, selected method, rejected alternative, feedback uptake, and unresolved trade-off.

The study has limitations. The linked Google Drive folder could be recognised as a shared qualitative-data location, but individual files were not directly retrievable in the present processing environment. The analysis therefore relied on the uploaded manuscript's qualitative evidence structure and accessible summaries, without inventing participant numbers, code frequencies, or verbatim quotations. A submission-stage version should insert selected anonymised quotations, specify the number and duration of interviews and observations, describe the institutional setting more fully, and archive an auditable ATLAS.ti project report. The findings are analytically transferable to similar Arabic translation-methodology courses, but they are not a statistical estimate of effects across institutions.

Conclusion

This goal-free evaluation examined what contextual systemic collaboration actually produced in Manahij al-Tarjamah rather than beginning with the programme's declared targets. ATLAS.ti-supported analysis identified a

Introduction to (Re-)Profiling T&I Education"; Ehrensberger-Dow, Delorme Benites, and Lehr, "A New Role for Translators and Trainers: MT Literacy Consultants."

connected pattern of benefits and constraints. Contextual cases redirected attention from isolated equivalence toward audience, genre, purpose, and cultural risk. Collaborative roles made translation reasoning visible and supported a professional metalanguage. Staged feedback and revision records strengthened evaluative judgment. Digital resources broadened alternatives and encouraged critical comparison, but also created verification dependency. Reflective portfolios supported transfer and evidence-based confidence, while excessive documentation created fatigue. Linguistic accuracy, participation equity, and the fidelity-readability dilemma remained uneven.

The overall judgment is conditional rather than celebratory. Contextual systemic collaboration optimised the quality of translation reasoning and revision more consistently than it optimised every feature of final-text accuracy. Its value lies in aligning case context, accountable roles, resource use, feedback, and reflection as one evidence chain. Future implementation should preserve that chain while simplifying documentation, calibrating terminology and grammar, auditing individual participation, and explicitly teaching machine-translation literacy. Goal-free evaluation contributes by ensuring that unexpected benefits and burdens remain visible in programme decisions.

References

- Bowker, Lynne. "Promoting Linguistic Diversity and Inclusion: Incorporating Machine Translation Literacy into Information Literacy Instruction for Undergraduate Students." *The International Journal of Information, Diversity, & Inclusion* 5, no. 3 (2021): 127–51. doi:10.33137/ijidi.v5i3.36159.
- Braun, V., and V. Clarke. "Using Thematic Analysis in Psychology." *Qualitative Research in Psychology* 3, no. 2 (2006): 77–101. doi:10.1191/1478088706qp0630a.
- Cahyanti, Aisyah Nur, and Abdul Malik. "Techniques and Quality Translation of Interrogative Sentence Dubbing: English-Indonesian and English-Arabic in The Cartoon Movie." *Arabiyat: Jurnal Pendidikan Bahasa Arab Dan Kebahasaaraban* 10, no. 2 (2023): 256–68. doi:10.15408/a.v10i2.34197.
- Carless, David, and David Boud. "The Development of Student Feedback Literacy: Enabling Uptake of Feedback." *Assessment & Evaluation in Higher Education* 43, no. 8 (2018): 1315–25. doi:10.1080/02602938.2018.1463354.

- Choiriyah, Nurul Aini, Muhammad Yunus Anis, Syed Nurulakla Syed Abdullah, and Nor Azuwan Yaakob. "Translation Quality of Arabic Directive Speech Acts in Pandemic's Book: An Attempt to Delve into Covid-19." *Arabiyat: Jurnal Pendidikan Bahasa Arab Dan Kebahasaaraban* 9, no. 1 (2022): 69–84. doi:10.15408/a.v9i1.25459.
- Colina, Sonia. "Translation Quality Evaluation: Empirical Evidence for a Functionalist Approach." *The Translator* 14, no. 1 (2008): 97–134. doi:10.1080/13556509.2008.10799251.
- Crescenzi-Lanna, Lucrezia. "Literature Review of Learner-Centered Design in Digital Learning Environments." *British Journal of Educational Technology* 53, no. 3 (2022): 716–35. doi:10.1111/bjet.13148.
- Daems, Joke, Sonia Vandepitte, Robert J. Hartsuiker, and Lieve Macken. "Translation Methods and Experience: A Comparative Analysis of Human Translation and Post-Editing with Students and Professional Translators." *Meta* 62, no. 2 (2017): 245–70. doi:10.7202/1041023ar.
- Defrancq, Bart. "Interprofessional Training for Student Conference Interpreters and Students of Political Science through Joint Mock Conferences: An Assessment." *The Interpreter and Translator Trainer* 16, no. 1 (2022). doi:10.1080/1750399X.2021.1919975.
- Dewi, H. D. "Translation Competences for Indonesian Novice Translators." *Journal of Language and Literature* 19, no. 1 (2019): 114–22. doi:10.24071/joll.v19i1.1817.
- Doherty, Stephen. "The Impact of Translation Technologies on the Process and Product of Translation." *International Journal of Communication* 10 (2016): 947–69.
- Edidarmo, Toto, and Zainal Muttaqin. "Foreignization Strategy in the Quranic Translation by MoRA: Study of Surah Al-Baqarah." *Arabiyat: Jurnal Pendidikan Bahasa Arab Dan Kebahasaaraban* 9, no. 1 (2022). doi:10.15408/a.v9i1.25207.
- Ehrensberger-Dow, Maureen, Alice Delorme Benites, and Caroline Lehr. "A New Role for Translators and Trainers: MT Literacy Consultants." *The Interpreter and Translator Trainer* 17, no. 3 (2023). doi:10.1080/1750399X.2023.2237328.
- Fereday, Jennifer, and Eimear Muir-Cochrane. "Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development." *International Journal of Qualitative Methods* 5, no. 1 (2006): 80–92. doi:10.1177/160940690600500107.

- Friese, Susanne. *Qualitative Data Analysis with ATLAS.Ti*. London: SAGE Publications, 2019.
- Godwin-Jones, Robert. "Partnering with AI: Intelligent Writing Assistance and Instructed Language Learning." *Language Learning & Technology* 26, no. 2 (2022): 5–24. doi:10.10125/73474.
- Hattie, J., and H. Timperley. "The Power of Feedback." *Review of Educational Research* 77, no. 1 (2007): 81–112. doi:10.3102/003465430298487.
- Herrington, Jan, and Ron Oliver. "An Instructional Design Framework for Authentic Learning Environments." *Educational Technology Research and Development* 48, no. 3 (2000): 23–48. doi:10.1007/BF02319856.
- Hmelo-Silver, Cindy E. "Problem-Based Learning: What and How Do Students Learn?" *Educational Psychology Review* 16, no. 3 (2004): 235–66. doi:10.1023/B:EDPR.0000034022.16470.f3.
- Hurtado Albir, Amparo. "The Acquisition of Translation Competence: Competences, Tasks, and Assessment in Translator Training." *Meta* 60, no. 2 (2015): 256–80. doi:10.7202/1032857ar.
- Johnson, David W., and Roger T. Johnson. "An Educational Psychology Success Story: Social Interdependence Theory and Cooperative Learning." *Educational Researcher* 38, no. 5 (2009): 365–79. doi:10.3102/0013189X09339057.
- Jonsson, Anders, and Gunilla Svingby. "The Use of Scoring Rubrics: Reliability, Validity and Educational Consequences." *Educational Research Review* 2, no. 2 (2007): 130–44. doi:10.1016/j.edurev.2007.05.002.
- Kasneci, Enkelejda, Kathrin Sessler, Stefan Küchemann, Maria Bannert, Daryna Dementieva, Frank Fischer, Urs Gasser, et al. "ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education." *Learning and Individual Differences* 103 (2023): 102274. doi:10.1016/j.lindif.2023.102274.
- Kelly, Dorothy. "Training the Trainers: Towards a Description of Translator Trainer Competence and Training Needs Analysis." *The Interpreter and Translator Trainer* 2, no. 1 (2008): 99–125. doi:10.1080/1750399X.2008.10798769.
- Kenny, Dorothy, and Stephen Doherty. "Statistical Machine Translation in the Translation Curriculum: Overcoming Obstacles and Empowering Translators." *The Interpreter and Translator Trainer* 8, no. 2 (2014): 276–94. doi:10.1080/1750399X.2014.936112.

- Kenny, John D., and Conor Cirkony. "Using Systems Perspectives to Develop Underlying Principles for Educational Reform." *Australian Journal of Teacher Education* 47, no. 1 (2022): 80–99. doi:10.14221/ajte.2022v47n1.6.
- Kiraly, Don. *A Social Constructivist Approach to Translator Education: Empowerment from Theory to Practice*. Manchester: St. Jerome Publishing, 2000.
- Kohnke, Lucas, Benjamin Luke Moorhouse, and Di Zou. "ChatGPT for Language Teaching and Learning." *RELC Journal* 54, no. 2 (2023): 537–50. doi:10.1177/00336882231162868.
- Kuswanto, Welly. "Implementation of English Learning at Dalwa Islamic Boarding School in The 5.0 Era For Generation Z." *Jurnal Pendidikan Islam* 13, no. 1 (2023): 93–109.
- Lee, Sang-Min. "The Effectiveness of Machine Translation in Foreign Language Education: A Systematic Review and Meta-Analysis." *Computer Assisted Language Learning* 36, nos. 1–2 (2023): 103–25. doi:10.1080/09588221.2021.1901744.
- Massey, Gary, Elsa Huertas-Barros, and David Katan. "Meeting Evolution with Innovation: An Introduction to (Re-)Profiling T&I Education." *The Interpreter and Translator Trainer* 17, no. 3 (2023). doi:10.1080/1750399X.2023.2237321.
- Mellinger, Christopher D. "Translators and Machine Translation: Knowledge and Skills Gaps in Translator Pedagogy." *The Interpreter and Translator Trainer* 11, no. 4 (2017): 280–93. doi:10.1080/1750399X.2017.1359760.
- Molina, Lucía, and Amparo Hurtado Albir. "Translation Techniques Revisited: A Dynamic and Functionalist Approach." *Meta* 47, no. 4 (2002): 498–512. doi:10.7202/008033ar.
- Moorkens, Joss. "What to Expect from Neural Machine Translation: A Practical In-Class Translation Evaluation Exercise." *The Interpreter and Translator Trainer* 12, no. 4 (2018): 375–87. doi:10.1080/1750399X.2018.1501639.
- Neves, Josélia. "Project-Based Learning for the Development of Social Transformative Competence in Socially Engaged Translators." *The Interpreter and Translator Trainer* 16, no. 4 (2022): 465–83. doi:10.1080/1750399X.2022.2084258.
- Nicol, David. "The Power of Internal Feedback: Exploiting Natural Comparison Processes." *Assessment & Evaluation in Higher Education* 46, no. 5 (2021): 756–78. doi:10.1080/02602938.2020.1823314.

- Niño, Ana. "Machine Translation in Foreign Language Learning: Language Learners' and Tutors' Perceptions of Its Advantages and Disadvantages." *ReCALL* 21, no. 2 (2009): 241–58. doi:10.1017/S0958344009000172.
- Nowell, L. S., J. M. Norris, D. E. White, and N. J. Moules. "Thematic Analysis: Striving to Meet the Trustworthiness Criteria." *International Journal of Qualitative Methods* 16, no. 1 (2017): 1–13. doi:10.1177/1609406917733847.
- O'Brien, Sharon. "Pauses as Indicators of Cognitive Effort in Post-Editing Machine Translation Output." *Across Languages and Cultures* 7, no. 1 (2006): 1–21. doi:10.1556/Acr.7.2006.1.1.
- Öner Bulut, Senem, and Nilüfer Alimen. "Translator Education as a Collaborative Quest for Insights into the Re-Positioning of the Human Translator (Educator) in the Age of Machine Translation: The Results of a Learning Experiment." *The Interpreter and Translator Trainer* 17, no. 3 (2023): 375–92. doi:10.1080/1750399X.2023.2237837.
- PACTE Group. "First Results of PACTE Group's Experimental Research on Translation Competence Acquisition: The Acquisition of Declarative Knowledge of Translation." *MonTI*, no. Special Issue 1 (2014): 85–115. doi:10.6035/MonTI.2014.ne1.2.
- Panadero, E., and A. Jonsson. "The Use of Scoring Rubrics for Formative Assessment Purposes Revisited: A Review." *Educational Research Review* 9 (2013): 129–44. doi:10.1016/j.edurev.2013.01.002.
- Patton, Michael Quinn. *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*. 4th ed. SAGE Publications, 2015. <https://us.sagepub.com/en-us/nam/qualitative-research-evaluation-methods/book232962>.
- Pym, Anthony. "Redefining Translation Competence in an Electronic Age: In Defence of a Minimalist Approach." *Meta* 48, no. 4 (2003): 481–97. doi:10.7202/008533ar.
- Risnawaty, Milisi Sembiring, Ihsan Fadilah, Henni Subagiarti, and Annim Hasibuan. "Macrostructure Analysis of Indonesian-Translation and Source Texts of the World Health Organization's Covid-19 Pandemic Guidance." *Indonesian Journal of Applied Linguistics* 12, no. 2 (2022): 533–44. doi:10.17509/ijal.v12i2.43858.
- Rogat, Toni K., Cindy E. Hmelo-Silver, Brittany H. Cheng, Alyson Traynor, Temitope Adeoye, Andrea Gomoll, and Brian Downing. "A Multidimensional Framework of Groups' Productive Disciplinary Engagement." *Frontline Learning Research* 10, no. 2 (2022): 1–21. doi:10.14786/flr.v10i2.874.

- Saleh, Asmalina, T. M. Phillips, Cindy E. Hmelo-Silver, Krista D. Glazewski, Bradford W. Mott, and James C. Lester. "A Learning Analytics Approach towards Understanding Collaborative Inquiry in a Problem-Based Learning Environment." *British Journal of Educational Technology* 53, no. 5 (2022): 1321–42. doi:10.1111/bjet.13198.
- Sarti, Gabriele, Arianna Bisazza, Ana Guerberofo-Arenas, and Antonio Toral. "DivEMT: Neural Machine Translation Post-Editing Effort across Typologically Diverse Languages." *Transactions of the Association for Computational Linguistics* 10 (2022): 1202–19. doi:10.1162/tacl_a_00518.
- Scriven, Michael. "Pros and Cons about Goal-Free Evaluation." *Evaluation Comment* 3, no. 4 (1972): 1–7.
- Stake, Robert E. *The Art of Case Study Research*. Thousand Oaks, CA: SAGE Publications, 1995.
- Stufflebeam, Daniel L. "The CIPP Model for Evaluation." In *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*, 279–317. Springer, 2000.
- Stufflebeam, Daniel L., and Chris L. S. Coryn. *Evaluation Theory, Models, and Applications*. San Francisco: Jossey-Bass, 2014.
- Terribile, Silvia. "Is Post-Editing Really Faster than Human Translation?" *Translation Spaces*, 2023. doi:10.1075/ts.22044.ter.
- Tracy, Sarah J. "Qualitative Quality: Eight Big-Tent Criteria for Excellent Qualitative Research." *Qualitative Inquiry* 16, no. 10 (2010): 837–51. doi:10.1177/1077800410383121.
- Woods, Megan, Rani Macklin, and Gemma K. Lewis. "Researcher Reflexivity: Exploring the Impacts of CAQDAS Use." *International Journal of Social Research Methodology* 19, no. 4 (2016): 385–403. doi:10.1080/13645579.2015.1023964.
- Youker, Brandon W., and Allyssa Ingraham. "Goal-Free Evaluation: An Orientation for Foundations' Evaluations." *The Foundation Review* 6, no. 4 (2014): 79–92. doi:10.9707/1944-5660.1229.