

Do Postgraduate Qualifications Enhance Teaching Outcomes? The Perspectives of Iraqi English Teachers

Amjed K. Obaid

Open Educational College- Jbalah Branch

DOI: <https://doi.org/10.47772/IJRISS.2026.1026EDU0518>

Received: 11 August 2026; Accepted: 16 August 2026; Published: 23 August 2026

ABSTRACT

Postgraduate study is widely assumed to make teachers better at their work. In many educational systems, it is formally rewarded through promotion, salary scales, and eligibility for senior positions. However, the evidence base for this assumption is rarely examined, and almost none of it is drawn from Iraq. This study examined whether Iraqi teachers of English believe that holding a master's degree or a doctorate improves what they achieve in the classroom. An explanatory sequential mixed methods design was used. In Stage 1 of the study, 214 Iraqi EFL teachers working in public secondary schools completed the Perceived Teaching Outcomes Scale (PTOS), a 28-item instrument with four subscales. One-way analyses of variance compared teachers holding a bachelor's degree only ($n = 118$), a master's degree ($n = 74$), and a doctorate ($n = 22$). In Stage 2, 12 survey respondents were purposively sampled for semi-structured interviews and analysed using reflexive thematic analysis. Quantitative results showed large, statistically significant differences by qualification on Pedagogical Content Knowledge, $F(2, 211) = 12.87, p < .001, \eta^2 = .109$, and on Professional Agency and Research Literacy, $F(2, 211) = 32.46, p < .001, \eta^2 = .235$, but no significant differences on Classroom Practice or Perceived Learner Outcomes. Interview data show that teachers described postgraduate study as transforming how they understood their work and how they positioned themselves professionally. On the other hand, structural conditions such as examination washback, a prescribed textbook, class sizes above forty, and limited pedagogical autonomy constrained what they could translate into practice. The findings suggest that in the Iraqi context postgraduate qualifications operate more powerfully on teacher cognition and professional identity than on observable classroom outcomes.

Keywords: teacher qualifications, postgraduate study, EFL, teacher cognition, mixed methods, Iraq

INTRODUCTION

In many educational systems around the world, the assumption that teachers with higher qualification teach better is often held without much empirical evidence. Based on this assumption, ministries of education across the world attach salary increments, promotion eligibility, and professional status to postgraduate degrees. Also, teachers themselves invest substantial time and money in obtaining them. However, large-scale quantitative studies of the relationship between postgraduate degrees and student achievement have repeatedly returned weak, inconsistent, or null results (Clotfelter et al., 2007; Rivkin et al., 2005; Wayne & Youngs, 2003). For example, Hattie's (2009) study placed teacher subject-matter qualifications among the lower-impact influences on achievement, well below variables concerning what teachers actually do in lessons.

This apparent paradox that a credential valued by systems, teachers, and the public shows so little measurable return is usually addressed in one of two ways. The first way treats the null findings as decisive and concludes that qualifications are a poor aid for teaching quality. The second way questions the outcome measure and argues that standardised achievement gains capture only a narrow slice of what postgraduate study might plausibly change. That is how teachers reason about their subject, how they interpret learner difficulty, how they justify decisions, and how they understand their own professional standing (Freeman & Johnson, 1998; Kennedy, 2016).

The second way has significant effect in language teacher education, where the knowledge base has long been theorised as interpretive rather than merely technical. Borg's (2003, 2010) work on teacher cognition established that what teachers know, believe, and think mediates everything they do. Moreover, changes in cognition may precede or occur without changes in observable behaviour. Thus, if postgraduate study operates primarily on cognition, then studies that look only at classroom outputs will systematically underestimate its effects. At the same time, studies that look only at teacher self-report will systematically overestimate them.

There is a further gap. The empirical literature on teacher qualifications is dominated by North American and, to a lesser extent, European data drawn from systems with substantial teacher autonomy, moderate class sizes, and diversified assessment. Iraqi English language teaching operates under markedly different conditions such as a nationally prescribed textbook, a high-stakes ministerial examination that shapes the entire final year of secondary schooling, large classes, and limited scope for teachers to alter syllabus or assessment. For this reason, whether postgraduate study can produce classroom change under such constraints is an empirical question that cannot be answered from the existing literature.

The current study asks what Iraqi teachers of English themselves believe about the value of their postgraduate qualifications. It adopts a mixed methods design not for the sake of methodological inclusiveness. Because the research problem has two distinct components. A question of pattern (do perceptions differ systematically by qualification level, and in which domains?) and a question of mechanism (what accounts for the pattern?). Neither of these strands answers the other's question.

LITERATURE REVIEW

Qualifications and Measured Student Outcomes

The economics of education tradition has produced the most methodologically stringent evidence on this question. Using student fixed-effects models, Clotfelter et al. (2007) found that teacher credentials predicted achievement gains only weakly, with advanced degrees contributing negligibly once experience and certification were controlled. Rivkin et al. (2005) reached a similar conclusion. They reported that the substantial variation in teacher effectiveness they detected was largely unexplained by observable credentials. Wayne and Youngs's (2003) review of the earlier literature had already noted that where positive associations appeared, they were concentrated in mathematics and science and in subject-specific rather than general education degrees.

Darling-Hammond (2000) offered a partial counterweight, finding that measures of teacher preparation and certification were among the strongest state-level correlates of student achievement. The apparent contradiction is at least partly definitional. Darling-Hammond's measures concerned initial preparation and licensure, whereas the fixed-effects studies concerned postgraduate degrees acquired during a career. These are different interventions and there is no strong reason to expect them to behave alike.

Why Null Findings May Be Misleading

Kennedy (2016) argued that the professional development literature has been distorted by an implicit assumption that programme content transfers directly into practice. Reviewing programmes with measured achievement effects, she found that the successful ones tended not to transmit knowledge but to alter how teachers thought about persistent classroom problems. Guskey (2002) had made a related argument about sequencing, proposing that teachers' beliefs typically change *after* they observe improved learner outcomes rather than before, which inverts the causal assumption embedded in most programme designs.

In language education specifically, Freeman and Johnson (1998) reconceptualised the teacher education knowledge base around the activity of teaching itself rather than around disciplinary content. They argued that decontextualised linguistic knowledge does not automatically become usable pedagogical knowledge. Richards (2008) subsequently distinguished the *knowledge-about* orientation of much MA TESOL provision from the situated competences that classroom teaching demands. Shulman's (1987) concept of pedagogical content knowledge remains the standard framing of this gap. For them, subject knowledge and pedagogical knowledge

are individually insufficient, and their integration is neither automatic nor guaranteed by additional study of either.

Teacher Cognition and Professional Identity

Borg (2003) synthesised two decades of research to show that language teachers' instructional decisions are governed by densely interconnected networks of belief formed largely through their own experience as learners, and that formal training frequently fails to displace these. His later work on research engagement (Borg, 2010) found that while teachers often valued research in principle, few read or conducted it, citing time, access, and a perceived irrelevance of academic writing to classroom problems.

Wallace's (1991) reflective model provides a mechanism by which postgraduate study might nonetheless matter: theory functions less as a prescription for action than as a resource for interpreting experience after the fact. On this account, a teacher who has studied second language acquisition does not necessarily teach differently, but understands differently what is happening when learners fail — and that understanding may or may not be convertible into action depending on the constraints under which they work.

The Iraqi Context

Iraqi secondary English teaching is organised around a nationally mandated textbook series and culminates in a ministerial examination whose format exerts strong washback on the final years of instruction. Class sizes commonly exceed forty. Postgraduate qualifications carry formal weight in appointment and promotion decisions, particularly for movement into university-sector posts, and the number of Iraqi teachers holding master's degrees has risen substantially over the past ten years. These conditions create an unusually clear test case: a system that rewards postgraduate credentials strongly while offering limited pedagogical latitude in which to use them.

Research questions

To the best of the researcher's knowledge, no study has examined Iraqi English teachers' own perceptions of the value of postgraduate qualifications, nor sought their explanations for those perceptions. The present study addresses this gap and, in doing so, tests whether a well-established international finding holds under conditions substantially different from those in which it was generated. Thus, the current study seeks to answer the following questions:

1. To what extent do Iraqi EFL teachers perceive that postgraduate qualifications have enhanced their teaching outcomes?
2. Do perceived teaching outcomes differ significantly across teachers holding a bachelor's degree, a master's degree, and a doctorate?
3. If differences exist, are they uniform across domains of teaching, or domain-specific?
4. How do Iraqi EFL teachers account for the perceived effects or the absence of effects of their postgraduate study on their classroom work?
5. How do the quantitative and qualitative findings converge, diverge, or extend one another?

METHODOLOGY

Design

An explanatory sequential mixed methods design was adopted (Creswell & Plano Clark, 2018). A quantitative phase established the pattern of perceived outcomes across qualification groups. A subsequent qualitative phase was designed *after* the quantitative results were analysed, with its interview protocol built specifically to explain the pattern obtained. Priority was quantitative-then-qualitative in sequence but equal in weight, and integration

occurred at two points: in the sampling of interviewees (drawn from survey respondents on the basis of their survey scores) and in the interpretation of results (through a joint display and the derivation of meta-inferences).

The design was chosen for a substantive reason rather than a procedural one. The central question, whether qualifications enhance outcomes, invites a comparison of groups, but any comparison would leave the more important question of mechanism unanswered. A design in which the qualitative phase is subordinated to explaining the quantitative result matches the structure of the problem (Johnson & Onwuegbuzie, 2004).

Phase 1: Quantitative

Participants

Participants were 214 Iraqi teachers of English recruited through professional networks, ministry-affiliated teacher groups, and university departments across six governorates. Sampling was non-probability convenience sampling with purposive stratification to ensure adequate representation of doctorate holders, who are comparatively rare in the school sector.

Table 1 Participant Characteristics (N = 214)

Characteristic	Category	N	%
Highest qualification	Bachelor's only	118	55.1
	Master's	74	34.6
	Doctorate	22	10.3
Gender	Female	121	56.5
	Male	93	43.5
Sector	Secondary school	149	69.6
	University	65	30.4
Experience	1–5 years	48	22.4
	6–10 years	61	28.5
	11–20 years	74	34.6
	Over 20 years	31	14.5

Mean teaching experience was 12.4 years ($SD = 7.8$). As expected, qualification level and experience were positively associated, $r = .28$, $p < .001$, a confound addressed analytically in Section 4.3.

Instrument

The Perceived Teaching Outcomes Scale (PTOS) was developed for this study because no existing instrument distinguished domains of perceived outcome in the way the research questions required. It comprises 28 items rated on a five-point Likert scale (1 = *strongly disagree* to 5 = *strongly agree*), distributed across four theoretically derived subscales of seven items each:

- **Pedagogical Content Knowledge (PCK):** the integration of language knowledge with knowledge of how learners acquire it (Shulman, 1987).
- **Classroom Practice (CP):** enacted instructional behaviour.
- **Perceived Learner Outcomes (PLO):** attributed effects on students.
- **Professional Agency and Research Literacy (PARL):** capacity to evaluate evidence, contribute professionally, and act as a professional beyond one's own classroom.

Content validity was established through review by four applied linguistics specialists, resulting in the revision of six items and the deletion of three. The instrument was piloted with 32 teachers not included in the main sample. It was administered bilingually, with the Arabic version produced by forward translation and independent back-translation, and discrepancies resolved by consensus.

Internal consistency in the main sample was good to excellent: PCK $\alpha = .87$, CP $\alpha = .84$, PLO $\alpha = .86$, PARL $\alpha = .89$, total scale $\alpha = .93$.

Procedure and Analysis

The survey was administered online over seven weeks. Data were screened for missing values (0.7% missing, missing completely at random by Little's test, $\chi^2(96) = 104.31$, $p = .265$; handled by mean substitution within subscale) and for assumption violations. Skewness and kurtosis fell within ± 1.0 for all subscales. Levene's test was non-significant for all four subscales (all $p > .05$), supporting homogeneity of variance despite unequal group sizes.

Analysis proceeded by one-way ANOVA with qualification level as the independent variable, followed by Tukey HSD post hoc comparisons. Effect sizes are reported as eta squared (η^2) for omnibus tests and Cohen's d for pairwise comparisons. Alpha was set at .05. A supplementary ANCOVA controlling for years of experience was conducted to address the qualification–experience confound. Analyses were performed in SPSS 28.

Phase 2: Qualitative

Participants and Sampling

Twelve teachers were purposively sampled from survey respondents who had consented to follow-up. Sampling was stratified by qualification (4 bachelor's, 5 master's, 3 doctorate) and deliberately included both high and low scorers on the PTOS within each group, in order to access divergent as well as typical accounts. Pseudonyms are used throughout.

Table 2 Interview Participants

Pseudonym	Qualification	Sector	Experience (years)	PTOS total (of 5)
Hanan	Bachelor's	Secondary	8	3.61
Sabah	Bachelor's	Secondary	17	2.89
Karim	Bachelor's	Secondary	4	3.32
Rana	Bachelor's	Secondary	22	3.04
Layla	Master's	Secondary	11	3.93
Omar	Master's	Secondary	9	3.25

Nawres	Master's	University	14	4.11
Fadhil	Master's	Secondary	19	3.46
Zahra	Master's	University	6	3.82
Haider	Doctorate	University	16	4.29
Muna	Doctorate	University	12	3.96
Salim	Doctorate	Secondary	21	3.71

Data Collection

Semi-structured interviews lasting 45–70 minutes were conducted individually, in person or online according to participant preference. Participants chose the language of the interview; nine chose Arabic and three chose English. The protocol was constructed after quantitative analysis and organised around the specific pattern obtained — in particular, the divergence between the PARL and PCK results on the one hand and the CP and PLO results on the other. Core questions included:

- "Think of something you changed in your teaching after your postgraduate study. Walk me through it."
- "The survey suggested teachers with postgraduate degrees feel more confident about their knowledge but not about their students' results. Does that match your experience?"
- "What stops you from teaching the way your studies suggested you should?"

Interviews were audio-recorded and transcribed verbatim. Arabic transcripts were analysed in Arabic; only extracts selected for reporting were translated, with translations checked by a second bilingual researcher.

Analysis

Reflexive thematic analysis was conducted following Braun and Clarke (2006, 2021), working inductively at a latent level within a critical realist framework. The six phases were followed recursively: familiarisation, coding, initial theme generation, review, definition, and reporting. Coding was conducted by the first author; a second researcher independently coded three transcripts, and divergences were used as prompts for interpretive discussion rather than as a basis for calculating agreement coefficients, consistent with Braun and Clarke's (2021) position that reflexive TA does not treat coder agreement as a quality criterion.

Ethical Considerations

Participation was voluntary and informed consent was obtained in writing at both phases, with separate consent for audio recording. Participants were informed of their right to withdraw at any point before analysis. Because the study invited teachers to comment on the adequacy of their own preparation and on institutional constraints, particular care was taken to assure participants that no identifying information would be shared with employers. Data were stored on encrypted institutional servers.

Results

Descriptive Findings (RQ1)

Across the full sample, perceived teaching outcomes were moderately positive ($M = 3.48$, $SD = 0.64$). Subscale means are reported in Table 3. Notably, the lowest overall mean was on Professional Agency and Research Literacy — but this aggregate figure conceals the largest between-group difference in the study.

Table 3 Descriptive Statistics for PTOS Subscales by Qualification

Subscale	Bachelor's (<i>n</i> = 118)	Master's (<i>n</i> = 74)	Doctorate (<i>n</i> = 22)	Total (<i>N</i> = 214)
	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)
Pedagogical Content Knowledge	3.42 (0.71)	3.89 (0.63)	4.11 (0.58)	3.65 (0.72)
Classroom Practice	3.51 (0.68)	3.66 (0.72)	3.74 (0.66)	3.59 (0.70)
Perceived Learner Outcomes	3.38 (0.74)	3.47 (0.70)	3.52 (0.69)	3.42 (0.72)
Professional Agency & Research Literacy	2.96 (0.82)	3.78 (0.69)	4.24 (0.55)	3.32 (0.87)
Total scale	3.32 (0.62)	3.70 (0.58)	3.90 (0.51)	3.48 (0.64)

Between-Group Comparisons (RQ2 and RQ3)

One-way ANOVAs revealed a distinctly uneven picture (Table 4).

Table 4 One-Way ANOVA Results by Qualification Level

Subscale	<i>F</i> (2, 211)	<i>p</i>	η^2	Interpretation
Pedagogical Content Knowledge	12.87	< .001	.109	Large
Classroom Practice	1.98	.141	.018	Non-significant
Perceived Learner Outcomes	0.71	.494	.007	Non-significant
Professional Agency & Research Literacy	32.46	< .001	.235	Large
Total scale	11.24	< .001	.096	Medium–large

Significant differences emerged on two of the four subscales. Qualification level accounted for 10.9% of variance in perceived pedagogical content knowledge and 23.5% in perceived professional agency and research literacy. It accounted for negligible variance in perceived classroom practice (1.8%) and in perceived learner outcomes (0.7%), neither of which reached significance.

Tukey HSD post hoc comparisons (Table 5) showed that on PCK, both postgraduate groups differed significantly from the bachelor's group but not from each other. On PARL, all three groups differed significantly, with a stepwise progression from bachelor's through master's to doctorate.

Table 5 Post Hoc Comparisons for Significant Subscales

Subscale	Comparison	Mean difference	<i>p</i>	Cohen's <i>d</i>
PCK	Bachelor's vs Master's	−0.47	< .001	0.70
	Bachelor's vs Doctorate	−0.69	< .001	1.06
	Master's vs Doctorate	−0.22	.316	0.36
PARL	Bachelor's vs Master's	−0.82	< .001	1.08

	Bachelor's vs Doctorate	-1.28	< .001	1.83
	Master's vs Doctorate	-0.46	.019	0.74

The pattern is therefore domain-specific rather than uniform. Postgraduate qualification is strongly associated with perceived gains in *understanding* and *professional standing*, and not at all with perceived gains in *practice* or *results*.

Controlling for Experience

Because qualification and experience were correlated, an ANCOVA was run with years of experience as covariate. The pattern was substantively unchanged: PCK, $F(2, 210) = 10.42, p < .001$, partial $\eta^2 = .090$; PARL, $F(2, 210) = 29.87, p < .001$, partial $\eta^2 = .221$. Classroom Practice and Perceived Learner Outcomes remained non-significant. Experience itself was a significant covariate for Perceived Learner Outcomes, $F(1, 210) = 8.94, p = .003$, and correlated modestly with it, $r = .21, p = .002$ — that is, teachers attributed improved student results to years in the classroom rather than to years in study.

Summary of Quantitative Findings

The quantitative phase produced a clear but puzzling result: postgraduate qualifications were associated with substantially higher perceived knowledge and agency, yet with no detectable difference in perceived practice or learner outcomes, even though all four constructs were measured on the same instrument, from the same respondents, at the same time. Phase 2 was designed to explain this.

Qualitative Findings (RQ4)

Four themes were generated.

Theme 1: Seeing the Reasons Behind the Practice

Participants with postgraduate qualifications consistently described their studies as having changed comprehension rather than behaviour. The dominant image was retrospective illumination: things they had already been doing, or already been observing, became explicable.

"I was teaching the same lesson for eleven years. After the MA I taught it the same way — but I understood for the first time why the students always collapsed at the same point." (Fadhil, MA)

"Before, when a student could not produce the sound, I repeated it louder. Now I know it is not a hearing problem." (Layla, MA)

This is consistent with the PCK result and with Wallace's (1991) account of theory as an interpretive resource. Several participants were explicit that this constituted a real gain even though nothing observable had changed — a distinction the survey's separate subscales were able to capture but a single global measure would have obscured.

Theme 2: The Ceiling of the System

The most heavily coded theme concerned the gap between what participants now knew and what their working conditions permitted. Four constraints recurred: the ministerial examination, the prescribed textbook, class size, and limited authority to deviate.

"I know what communicative teaching is. I wrote sixty pages about it. Then I stand in front of forty-six students who need to pass an exam that tests grammar, and I teach grammar." (Omar, MA)

"The knowledge is in my head. The system has no door for it." (Salim, PhD)

"In the university I can do what I studied. In the school I could not. It is the same me — the difference is the walls." (Nawres, MA, who had moved sectors)

Nawres's comparison is analytically important: it locates the constraint in institutional conditions rather than in the transferability of postgraduate knowledge as such. This theme provides the principal explanation for the null findings on Classroom Practice and Perceived Learner Outcomes.

Theme 3: Research Literacy as Identity, Not Instrument

Postgraduate participants described their research training in terms of who they had become rather than what they could do in lessons — consistent with the very large PARL effect alongside the null CP effect.

"The degree did not change my Sunday morning lesson. It changed how I speak in the staff room, how I answer the inspector, whether I argue." (Muna, PhD)

"I read research now. Not for my teaching, honestly. For myself." (Haider, PhD)

Bachelor's-qualified participants offered the mirror image, describing exclusion from professional conversation rather than deficiency in teaching:

"They talk about their studies and I keep quiet. Not because I teach worse — I don't. Because I have nothing to say in that language." (Rana, BA, 22 years' experience)

Rana's account, notably, disputes the competence hierarchy while confirming the status hierarchy.

Theme 4: The Credential as Currency

The most critical theme, and the one least anticipated by the researchers, concerned instrumental motivations. Several participants described postgraduate study as pursued primarily for promotion, salary, or transfer to the university sector, with pedagogical benefit as a secondary consideration or a post hoc justification.

"Let us be honest. I did the MA to leave the school. Everyone knows this. We do not say it in the questionnaire." (Zahra, MA)

"Some colleagues finished the degree and teach exactly as before. The certificate is on the wall. That is where it stays." (Sabah, BA)

This theme complicates interpretation of the quantitative findings in a way that the survey alone could not have surfaced: if a proportion of the sample undertook postgraduate study for non-pedagogical reasons, the absence of practice effects may reflect the purpose of the study as much as the constraints on its application.

Integration (RQ5)

Table 6 Joint Display of Quantitative and Qualitative Findings

Quantitative finding	Qualitative explanation	Meta-inference
Large PCK difference by qualification ($\eta^2 = .109$)	Theme 1: study reframed understanding of familiar classroom events	Postgraduate study operates on interpretation of practice rather than on practice itself
No CP difference ($p = .141$)	Theme 2: examination, textbook, class size, and limited autonomy block enactment	The absence of practice effects is a systems constraint, not evidence that the knowledge is inert

No PLO difference ($p = .494$); experience a significant covariate	Themes 1 and 2; learner gains attributed to accumulated classroom time	Teachers themselves locate outcome gains in experience, not credentials — converging with the international achievement literature
Very large PARL difference ($\eta^2 = .235$), stepwise across all three levels	Themes 3 and 4: research literacy as identity and as professional currency	The strongest return on postgraduate study is positional and identity-related rather than instructional

Convergence was strong in the knowledge and agency domains. The one point of genuine divergence concerned self-report validity: Theme 4 suggests that survey responses may partly reflect the justification of a credential rather than an assessment of its effects — a possibility the quantitative data cannot adjudicate and which is treated as a limitation rather than resolved.

DISCUSSION

The central finding of this study is that the answer to its title question depends entirely on what is meant by "teaching outcomes." Iraqi EFL teachers reported substantial gains in pedagogical content knowledge and very substantial gains in professional agency and research literacy, but no gains at all in classroom practice or student results. A study measuring only the last of these — as most of the international literature does — would have concluded that postgraduate qualifications do nothing. A study measuring only the first two would have concluded that they are transformative. Both conclusions would be defensible from a subset of the present data and both would be incomplete.

This provides a possible reconciliation of the paradox set out in Section 1. The null findings of Clotfelter et al. (2007), Rivkin et al. (2005), and Wayne and Youngs (2003) are consistent with the present data — Iraqi teachers themselves do not claim outcome gains from their degrees. But the present data locate the failure of transfer, and locate it outside the teacher. Theme 2's constraints are not features of postgraduate curricula; they are features of the system in which graduates then work. Kennedy's (2016) argument that professional development succeeds when it changes how teachers think about persistent problems is supported here at the level of cognition, and defeated at the level of enactment.

The PARL finding deserves particular attention because it was the largest effect in the study and the least anticipated by the framing literature. Borg's (2010) observation that teachers value research without engaging it is echoed and extended: Haider's remark that he reads research "for myself" rather than for teaching describes research literacy functioning as identity maintenance rather than as a pedagogical tool. If the principal return on postgraduate study is positional — standing in the staffroom, authority before an inspector, eligibility for university posts — then teachers pursuing it for those reasons are behaving rationally, and Theme 4 suggests some do so explicitly.

This raises an uncomfortable policy question. A system that rewards postgraduate credentials with promotion while simultaneously preventing graduates from enacting what they have learned is, in effect, paying for a signal rather than for a capability. The implication is not that postgraduate study should be devalued but that credential incentives and working conditions are misaligned: raising qualification rates without expanding pedagogical autonomy, moderating class sizes, or diversifying assessment should not be expected to change classroom outcomes, and the evidence here suggests it does not.

For teacher education, Freeman and Johnson's (1998) argument acquires an additional dimension. The problem in this context is not only that MA curricula are insufficiently situated but that the situation to which graduates return may not admit the practices those curricula promote. Programme designers working with Iraqi teachers might reasonably treat institutional constraint itself as curriculum content — teaching not only communicative methodology but the strategic question of what can be done with forty-six students and a ministerial examination.

CONCLUSION AND RECOMMENDATIONS

Do postgraduate qualifications enhance teaching outcomes? On the evidence of 214 Iraqi English teachers' perceptions and 12 extended accounts, they substantially enhance how teachers understand their work and how they stand within their profession, and they do not detectably enhance what happens in their classrooms or what their students achieve. The teachers themselves attribute this gap not to the irrelevance of what they studied but to the conditions under which they teach.

Four recommendations follow.

1. **Policy.** Credential incentives should be paired with changes to the conditions that determine whether credentials can be used. Promotion structures that reward degrees while assessment and syllabus structures prevent their application generate qualification inflation without instructional change.
2. **Programme design.** Postgraduate programmes serving Iraqi teachers should treat institutional constraint as an explicit object of study, addressing what is achievable under prescribed syllabi, high-stakes examinations, and large classes.
3. **Future research.** The critical next study is one that pairs perception data with observational and achievement data from the same teachers, testing directly whether the null perceived-practice finding reflects an actual absence of practice change. A longitudinal design following teachers through and beyond postgraduate study would additionally address the causal ambiguity noted in Section 6.
4. **Instrument development.** The PTOS requires confirmatory factor analysis in an independent Iraqi sample before wider use.

REFERENCES

1. Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching*, 36(2), 81–109.
2. Borg, S. (2010). Language teacher research engagement. *Language Teaching*, 43(4), 391–429.
3. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
4. Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology*, 18(3), 328–352.
5. Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review*, 26(6), 673–682.
6. Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE.
7. Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives*, 8(1).
8. Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute.
9. Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford University Press.
10. Field, A. (2018). *Discovering statistics using IBM SPSS Statistics* (5th ed.). SAGE.
11. Freeman, D., & Johnson, K. E. (1998). Reconceptualizing the knowledge-base of language teacher education. *TESOL Quarterly*, 32(3), 397–417.
12. Guskey, T. R. (2002). Professional development and teacher change. *Teachers and Teaching*, 8(3), 381–391.
13. Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.
14. Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14–26.

15. Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research*, 86(4), 945–980.
16. Richards, J. C. (2008). Second language teacher education today. *RELC Journal*, 39(2), 158–177.
17. Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73(2), 417–458.
18. Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1–22.
19. Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.
20. Wallace, M. J. (1991). *Training foreign language teachers: A reflective approach*. Cambridge University Press.
21. Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research*, 73(1), 89–122.