GLODAL

OPEN ACCESS

*Brief Reports in Language Research*

*Brief Reports in Language Research*, 1, 260105 (2026)
https://doi.org/10.65553/BRLR.260105

# *Which Vocabulary Teacher Competencies do Experts Agree on? A fuzzy Delphi Evaluation of the Vocabulary Teachers' Competencies Framework*

Wei Lun Wong*[1] Lawrence Jun Zhang[2]

[1]*Faculty of Education, Universiti Kebangsaan Malaysia, Malaysia*
[2]*Faculty of Arts and Education, The University of Auckland, Auckland, New Zealand*

Abstract

The Vocabulary Teachers' Competencies (VTC) framework specifies 18 competencies across three domains. But it was proposed without structured expert validation. This brief report presents a single round fuzzy Delphi evaluation of the framework. Twenty-one experts rated 55 indicators that operationalise the 18 competencies on a 7-point importance scale. Ratings were converted to triangular fuzzy numbers, aggregated and defuzzified, and consensus was assessed with the coefficient of variation (CV) and Kendall's W. Experts rated the indicators as important overall (M = 5.61). However, only four indicators reached the defuzzified threshold of G ≥ 5.0, all in Domains 1 and 2. Eleven indicators achieved high consensus (CV < .15), 41 moderate consensus and three low consensus (CV ≥ .30). All three low-consensus indicators operationalised Competency 3.4, cross-curricular collaboration. Domain 3 (Professional and Socio-Cultural Capabilities) received the lowest and most dispersed ratings. Overall concordance was significant but weak (W = .137, $p$ < .001). The findings support the knowledge and skills domains as a stable core and suggest that competencies dependent on institutional opportunity need redefinition before the VTC is used for assessment.

*Keywords:* vocabulary teacher competencies; fuzzy Delphi method; expert consensus; teacher education; content validation

Submission date: August 30, 2026; Revision date: November 3, 2026; Online: November 14, 2026

Copyright ©2026
The authors

(corresponding author: colinw@ukm.edu.my)

## Introduction

### *Why English Teachers Need a Shared Framework*

Every English teacher knows that words matter. Students who know more words read more easily, understand more of what they read and find it easier to learn new subject content. Vocabulary research has also produced clear guidance for teaching. Beck et al. (2013) show how to choose words by sorting them into tiers. Also, Nation (2022) shows how to balance vocabulary learning across four strands, from meaning-focused reading and listening to deliberate study and fluency practice. However, many English teachers report that they were never shown how to teach vocabulary well. Green and Coxhead (2026) argue that teacher education has lacked a shared description of what an English teacher needs to know and be able to do. Their answer is the Vocabulary Teachers' Competencies (VTC) framework. It sets out 18 competencies in three domains, with six in each:

1. Knowledge of vocabulary research and resources: Understanding how words are built and the different types of vocabulary students meet.

2. Vocabulary analysis and instructional skills: Choosing useful words to teach and judging the vocabulary demands of a text.

3. Professional and socio-cultural capabilities: Planning over time, keeping up with research, collaborating across subjects, differentiating and teaching inclusively.

The domains follow the classic view of competence as knowledge, skills and dispositions. Richey et al. (2001, p. 8) define a competency as "a knowledge, skill, or attitude that enables one to effectively perform the activities of a given occupation or function to the standards expected in employment."

### *Are all Competencies the Same?*

Research on teacher knowledge suggests that they are not. Shulman (1986) showed that knowing a subject is not the same as knowing how to make it learnable for students. Teachers need pedagogical content knowledge that joins the two. Blömeke et al. (2015) describe competence as a continuum, running from knowledge and dispositions, through what teachers notice and decide in a situation, to what they actually do in class. Opfer and Pedder (2011) add that teacher learning is a complex system shaped by schools, colleagues and policy, not a simple path from knowledge to practice. Together, these views suggest that some competencies can be learned by an individual, while others depend heavily on context. Teng and Wong (2026) apply this thinking to the VTC. Following "Sam", a Year 6 teacher preparing a science lesson on erosion, they show how the competencies must work together in a real week of teaching. They also name "frictions" that teacher educators may meet in practice. Three matter here:

- Uneven size: Some competencies describe one ability. Others, especially in Domain 3, bundle several together.

- Opportunity, not only ability: Collaboration and keeping up with research depend on timetables, leadership and access that a single teacher may not control.

- Equity placed last: Listing equity and inclusion at the end may suggest it is an add-on, when in practice it shapes decisions from the start.

Wong et al. (2026) examined how the three VTC domains could be measured through teacher self-report. They developed the VTC Questionnaire and evaluated it with 267 participants from a predominantly Malaysian sample. Participants reported high levels of perceived competence across all three domains. The strict three-factor confirmatory model showed marginal fit. The

three-factor exploratory structural equation model provided the best fit, suggesting that the domains were distinguishable and closely related. Scores showed strong reliability. Evidence for convergent and discriminant validity was mixed, and measurement invariance across pre-service and in-service English teachers received approximate support. These findings describe how teachers' perceptions were structured across the three domains. A complementary question is which indicators experts judge important and whether the degree of agreement varies by domain.

### *The gap*

VTC was built from an integrative review and refined by its authors and peers. Green and Coxhead (2026, p. 5) decided that "conducting a new empirical study (e.g., meta-analysis, Delphi consensus etc.) was deemed unnecessary and methodologically excessive." Hence, Teng and Wong (2026) describe VTC as "an expert-informed content framework with preliminary content validity." Before a framework is used to plan courses or assess teachers, we need evidence that supports those uses (Kane, 2013). The first piece of such evidence is whether experts agree on its content. No study has yet provided it. It is also unknown whether experts judge the three domains equally, or whether, they agree less about the context-dependent competencies of Domain 3.

### *This study*

A Delphi study is a standard way to gather this evidence. A panel of experts rates a set of statements independently, so that no single strong voice dominates and researchers examine how far they agree (Hsu & Sandford, 2007). The fuzzy Delphi method (FDM) treats each rating as a small range rather than an exact point, which better reflects how people judge and allows screening in fewer rounds (Ishikawa et al., 1993; Murray et al., 1985). Agreement can be checked both for each item and for the panel's overall ranking of items (Schmidt, 1997).

This brief report makes three contributions. It offers the first empirical expert evaluation of VTC. It identifies which indicators form an agreed core for teacher education. Also, it tests whether expert agreement differs across the three domains, providing evidence on the concerns raised by Teng and Wong (2026). We ask:

**RQ1.** Which VTC indicators do experts see as essential?

**RQ2.** How far do experts agree, and does their agreement differ across the three domains?

## Methods

### *Participants*

The fuzzy Delphi panel comprised 21 experts. The research team identified potential panellists through its professional networks and by reviewing relevant publications, official institutional webpages and LinkedIn profiles. The stated eligibility criteria were a permanent academic appointment at the rank of lecturer or above, expertise in TESL or TESOL and in vocabulary research, and at least five years of teaching or research experience.

The panel included 14 men and seven women aged 29–64 years (M = 42.48, SD = 9.91). All held doctoral degrees. Their roles included professors, lecturers, researchers, postdoctoral researchers and a language education specialist. Across all 21 panelists, teaching or research experience ranged from 5 to 39 years (M = 16.10, SD = 9.64).

The 21 interpretable geographical responses covered 13 national or regional contexts across East, Southeast, South and West Asia. The most frequently reported areas of expertise were

applied linguistics (n = 19), TESOL (n = 15) and teacher education (n = 14). Three panellists selected vocabulary studies. Panellists could report multiple areas of expertise. The requirement for ethics approval was waived because the study was deemed to pose no risk to participants. Panelists provided informed consent by ticking a required checkbox in Google Forms.

### *Instrument*

We developed the questionnaire to operationalise the 18 VTC competencies (Green & Coxhead, 2026). Before the Delphi round, experts reviewed the draft questionnaire for layout, wording and item clarity. Their feedback informed two rounds of revision. To address double-barrelled wording, we split items that combined distinct ideas or practices into separate statements so that panellists could rate each one independently.

The final questionnaire comprised 55 indicators. Competency 1.1 had four indicator and each of the other competencies had three. We coded each indicator by domain, competency and item number. For example, 3.4.2 identified the second indicator for Competency 3.4 in Domain 3. During the Delphi round, panelists rated the importance of each indicator on a 7-point scale ranging from 1 (*not important at all*) to 7 (*extremely important*).

### *Procedure and Analysis*

The data come from a single Delphi round in February 2026 using Google Form. We analysed them in three steps, each answering a simple question.

**Step 1: How important is each indicator? (fuzzy screening)**

Each rating x was turned into a fuzzy number with three points, (l, m, u) = (max(x − 2, 0), x − 1, x). For example, a rating of 6 became (4, 5, 6). This left-shifted version treats the rating as the top of the expert's range and so gives a cautious estimate. The fuzzy numbers were then combined across all 21 experts: L is the lowest l given by any expert, M is the average m, and U is the highest u. Finally, the three values were averaged into one score:

$$G = \frac{L + M + U}{3}$$

An indicator was judged essential when $G \geq 5.0$. Because L comes from the lowest rating, one expert who rates an indicator low pulls its G score down. This is deliberate: an indicator counts as essential only when the whole panel supports it.

**Step 2: Do experts agree on each indicator? (consensus)**

We used the coefficient of variation (CV), the standard deviation divided by the mean. It shows how spread out the ratings are relative to their average, so a small CV means that experts gave similar ratings. We used three bands: high ($CV < .15$), moderate ($.15 \leq CV < .30$) and low ($CV \geq .30$).

**Step 3: Do experts rank the indicators in a similar order? (concordance)**

Kendall's W runs from 0 (no shared rank order) to 1 (identical rank orders). Schmidt (1997) suggests reading values of about .1 as very weak, .3 as weak, .5 as moderate and .7 as strong agreement. W asks a different question from CV: not whether experts gave one indicator similar scores, but whether they placed all 55 indicators in a similar order of importance.

Decisions combined importance and agreement. Indicators with $G \geq 5.0$ and high consensus were classed as core. Other high-consensus indicators were retained. Moderate-consensus indicators were flagged for revision, and low-consensus indicators for deletion or reconceptualisation.

## Results

### *Importance (RQ1)*

Experts rated the indicators as important overall (M = 5.61 on the 7-point scale; item means 4.76–6.33). Defuzzified values were lower (G = 3.59–5.11; mean 4.33). Only four of 55 indicators (7.3%) met G ≥ 5.0: 1.2.3 and 1.3.3 (G = 5.11), 1.5.2 (5.08) and 2.5.2 (5.00). Because U = 7 for every indicator, G depended on L, set by the lowest single rating, and on M. In practice, G ≥ 5.0 required that no expert rated an indicator below 5 and that its mean rating reached about 6.0.

Importance declined across domains (Table 1). Domain 1 had the highest mean rating (5.83) and three core indicators, Domain 2 had one and Domain 3 had none. The highest Domain 3 value was G = 4.68 (3.5.3). At competency level (Table 2), the lowest mean G values were for 3.4, collaboration (3.60); 1.4, sourcing and evaluating resources (3.85); and 3.2, engaging with new research (3.89).

**Table 1.** *Importance and consensus by VTC domain*

| Domain | Indicators | Mean rating (range) | Mean G (range) | Mean CV | High / moderate / low consensus | G ≥ 5.0 |
|---|---|---|---|---|---|---|
| **1. Knowledge of vocabulary research and resources** | 19 | 5.83 (5.33–6.33) | 4.47 (3.79–5.11) | .181 | 5 / 14 / 0 | 3 |
| **2. Vocabulary analysis and instructional skills** | 18 | 5.65 (5.43–6.00) | 4.46 (3.81–5.00) | .169 | 6 / 12 / 0 | 1 |
| **3. Professional and socio-cultural capabilities** | 18 | 5.33 (4.76–6.05) | 4.05 (3.59–4.68) | .233 | 0 / 15 / 3 | 0 |
| **Total** | 55 | 5.61 (4.76–6.33) | 4.33 (3.59–5.11) | .194 | 11 / 41 / 3 | 4 |

*Note*. N = 21 experts. Ratings on a 7-point scale. G = defuzzified value of the aggregated triangular fuzzy number.

### *Consensus and concordance (RQ2)*

Eleven indicators reached high consensus, 41 moderate and three low (Table 3). High-consensus indicators came only from Domains 1 (n = 5) and 2 (n = 6). Mean CV was lowest in Domain 2 (.169) and highest in Domain 3 (.233). All three low-consensus indicators operationalised Competency 3.4, collaborative partnerships across the curriculum (CV = .346–.387). They also had the lowest means (4.76–4.86), the largest SDs (1.68–1.84) and minimum ratings of 1. Two further Domain 3 indicators approached the low-consensus threshold: 3.2.3 (CV = .297) and 3.6.1 (CV = .282). Indicator 1.1.2 lay on the high/moderate boundary (CV =

.150) and was classed as moderate.

High consensus did not guarantee high importance. Seven high-consensus indicators fell short of the G threshold, including 1.2.1 (CV = .137, G = 4.90) and 2.2.2 (CV = .146, G = 4.95). Kendall's W was significant but small, W = .137, $\chi^2(54) = 155.23$, $p < .001$. Experts' rank orderings were not random, but their agreement on the relative importance of indicators fell between Schmidt's (1997) benchmarks for very weak (.1) and weak (.3) agreement. Applying the decision rules gave 4 core, 7 retained, 41 to revise and 3 to delete or reconceptualise.

**Table 2.** *Results by VTC competency*

| Competency | Mean rating | Mean G (range) | Mean CV | Core / retained / revise / delete |
|---|---|---|---|---|
| **1.1 Statistical nature of language** | 5.91 | 4.47 (3.81–4.75) | .186 | 0 / 0 / 4 / 0 |
| **1.2 Structures of vocabulary** | 6.00 | 4.89 (4.65–5.11) | .139 | 1 / 1 / 1 / 0 |
| **1.3 Types of vocabulary** | 5.82 | 4.50 (4.11–5.11) | .181 | 1 / 0 / 2 / 0 |
| **1.4 Sourcing and evaluating resources** | 5.55 | 3.85 (3.79–3.92) | .237 | 0 / 0 / 3 / 0 |
| **1.5 Research-supported pedagogies** | 5.68 | 4.56 (4.16–5.08) | .176 | 1 / 0 / 2 / 0 |
| **1.6 Theories of vocabulary learning** | 6.02 | 4.56 (4.27–4.76) | .164 | 0 / 1 / 2 / 0 |
| **2.1 Identifying word features** | 5.67 | 4.45 (4.21–4.59) | .177 | 0 / 0 / 3 / 0 |
| **2.2 High-utility instructional words** | 5.62 | 4.54 (4.14–4.95) | .154 | 0 / 2 / 1 / 0 |
| **2.3 Components of word knowledge** | 5.66 | 4.44 (4.19–4.63) | .168 | 0 / 2 / 1 / 0 |
| **2.4 Atypical vocabulary development** | 5.70 | 4.57 (4.56–4.57) | .162 | 0 / 1 / 2 / 0 |
| **2.5 Vocabulary demands of texts** | 5.75 | 4.36 (3.81–5.00) | .180 | 1 / 0 / 2 / 0 |
| **2.6 Relationships between words** | 5.51 | 4.39 (4.17–4.51) | .175 | 0 / 0 / 3 / 0 |
| **3.1 Structured approach to instruction** | 5.29 | 3.98 (3.73–4.14) | .215 | 0 / 0 / 3 / 0 |
| **3.2 Engaging with new research** | 5.33 | 3.89 (3.67–4.19) | .240 | 0 / 0 / 3 / 0 |
| **3.3 Interest, assessment and feedback** | 5.38 | 4.24 (4.08–4.48) | .180 | 0 / 0 / 3 / 0 |
| **3.4 Cross-curricular collaboration** | 4.81 | 3.60 (3.59–3.62) | .364 | 0 / 0 / 0 / 3 |
| **3.5 Differentiated instruction** | 5.75 | 4.47 (4.19–4.68) | .176 | 0 / 0 / 3 / 0 |
| **3.6 Equity and inclusion** | 5.41 | 4.14 (3.65–4.54) | .225 | 0 / 0 / 3 / 0 |

*Note*. Competency labels abridged from Green and Coxhead (2026). Values are means across each competency's indicators.

**Table 3.** *Indicators with high or low consensus*

| Indicator | Mean | SD | G | CV | Decision |
|---|---|---|---|---|---|
| **1.2.3** | 6.33 | 0.80 | 5.11 | .126 | Core |
| **1.3.3** | 6.33 | 0.80 | 5.11 | .126 | Core |
| **1.5.2** | 6.24 | 0.83 | 5.08 | .133 | Core |
| **2.5.2** | 6.00 | 0.84 | 5.00 | .139 | Core |
| **2.2.2** | 5.86 | 0.85 | 4.95 | .146 | Retained |
| **1.2.1** | 5.71 | 0.78 | 4.90 | .137 | Retained |
| **1.6.3** | 6.29 | 0.90 | 4.76 | .144 | Retained |
| **2.3.1** | 5.90 | 0.83 | 4.63 | .141 | Retained |
| **2.4.1** | 5.71 | 0.85 | 4.57 | .148 | Retained |
| **2.2.3** | 5.57 | 0.81 | 4.52 | .146 | Retained |
| **2.3.3** | 5.52 | 0.81 | 4.51 | .147 | Retained |
| **3.4.1** | 4.86 | 1.68 | 3.62 | .346 | Delete or reconceptualise |
| **3.4.2** | 4.81 | 1.72 | 3.60 | .358 | Delete or reconceptualise |
| **3.4.3** | 4.76 | 1.84 | 3.59 | .387 | Delete or reconceptualise |

*Note*. Core = G ≥ 5.0 and CV < .15; retained = CV < .15 and G < 5.0; delete or reconceptualise = CV ≥ .30. The remaining 41 indicators (moderate consensus) were flagged for revision; full results are available as supplementary material.

## Discussion and conclusion

What do these results mean for teachers, student teachers and the people who train them? We see two main messages.

### *Experts agree on what vocabulary teachers need to know and do*

All four core indicators and all eleven high-consensus indicators came from Domains 1 and 2. The core indicators sit under four competencies: understanding the structures of vocabulary (1.2), understanding different types of vocabulary (1.3), knowing research-supported ways to teach vocabulary (1.5), and identifying the vocabulary demands of texts (2.5).

These are long-standing themes in vocabulary research. Beck et al. (2013) sort words into three tiers: everyday words that most students pick up naturally (Tier 1), useful words that appear across many texts and subjects, such as *analyse or consequence* (Tier 2) and subject-specific words such as *photosynthesis* (Tier 3). They recommend that teaching focus on Tier 2 words and knowing about such word types is what Competency 1.3 describes. Nation (2022) explains why the vocabulary demands of a text matter. Readers need to know a very high proportion of the words in a text to understand it comfortably. Hence, teachers need to judge whether a text is within reach and which words to teach before reading. Nation (2022) also recommends balancing vocabulary learning across four strands: learning from reading and listening, using words in speaking and writing, deliberate study of words and fluency practice.

For teachers, this is reassuring. The panel strongly supported what vocabulary research has said for years. The result supports Green and Coxhead's (2026) view that the field already shares a mature consensus, at least for these two domains.

### *Experts disagree about the professional and social side*

Domain 3 told a different story. It had no core or high-consensus indicators, the lowest importance scores and all three low-consensus indicators. The clearest case was collaboration

across the curriculum (3.4). Some experts rated these indicators as highly important, while others gave them a 1.

Teng and Wong's (2026) idea of "opportunity, not only ability" offers one explanation. In their example, Sam's Year 6 team agrees to teach both the everyday and the science meanings of words such as *deposit and transport* across three subjects. That only works because the team has a shared planning meeting. As Teng and Wong (2026, p. 8) put it, whether Sam's practice improves "depends on the timetable as much as on Sam." Experts who see competencies as abilities an individual teacher can be trained in may hesitate to include collaboration, because it relies on school structures. Experts who see vocabulary teaching as a whole-school effort may rate it very highly. Therefore, the disagreement may be less about whether collaboration matters than about whether it belongs on a list of individual competencies.

The same reading fits other Domain 3 results. Engaging with new research (3.2.3) and equity and inclusion (3.6.1) came close to the low-consensus threshold. Both depend partly on time, access and school culture, and both pack broad ideas into a single competency. This matches research that treats teacher learning as a complex system shaped by schools and policy, not as a simple path from knowledge to practice (Opfer & Pedder, 2011). Even in Domain 1, sourcing and evaluating resources (1.4) received low importance scores, perhaps because access to good materials varies widely between schools.

### *Implications for theory*

The three VTC domains do not have equal empirical standing. Domains 1 and 2 behave like a stable core, whereas Domain 3 behaves like a context-dependent layer. This supports Teng and Wong's (2026) call to separate what a teacher is able to do, what they actually do in class and the conditions that make it possible. It also fits Blömeke et al.'s (2015) view of competence as a continuum. Knowledge and dispositions become effective teaching only through what teachers notice, interpret and decide in real situations. Shulman's (1986) idea of pedagogical content knowledge makes a related point. Knowing about words is not the same as knowing how to make them learnable for a particular class.

### *Implications for Teachers and Teacher Educators*

Start with the core: The four core indicators offer a defensible minimum for any vocabulary course or professional development session: how words are built, what kinds of words students meet, which teaching approaches have research support and how to judge a text's vocabulary load.

Treat Domain 3 as shared goals, not individual scores: Collaboration, research engagement and equity are better framed as context-sensitive goals for schools and teams than as boxes on one teacher's checklist.

Redefine collaboration: Before the VTC is used for self-assessment or evaluation, Competency 3.4 should separate a teacher's own collaborative knowledge and skills from the school conditions that make collaboration possible.

### *Limitations*

This study has five limitations. First, the findings come from a single round. A second round that shows experts the group results and gives clearer definitions for Domain 3 would show whether the disagreement is about wording or substance. Second, the fuzzy method was cautious. Left-shifted fuzzy numbers and the use of the lowest and highest ratings make G sensitive to one low rating, and U was 7 for every indicator. Other common choices, such as symmetric fuzzy numbers or geometric-mean aggregation (Ishikawa et al., 1993) might change

which indicators are kept. Third, the G ≥ 5.0 cut-off was strict. Lowering it to 4.5 would keep 24 indicators (10, 11 and 3 across Domains 1–3). Fourth, most ratings fell between 5 and 7. This ceiling limits how differently experts can rank the indicators and partly explains the small W (Schmidt, 1997). Fifth, panel composition limits how far the results generalise. Expert ratings of importance are also only one kind of evidence. They do not show how the indicators work in assessment or whether they relate to student learning (Kane, 2013).

### *Conclusion*

Experts broadly agree on what English teachers need to know and be able to do, and the VTC's first two domains capture this well. The professional and socio-cultural domain and collaboration needs clearer definitions and a further Delphi round before it is used to assess teachers.

## References


Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life: Robust vocabulary instruction* (2nd ed.). Guilford Press.

Blömeke, S., Gustafsson, J.-E., & Shavelson, R. J. (2015). Beyond dichotomies: Competence viewed as a continuum. *Assessment of Competiences in Higher Education, 223*(1), 3–13. https://doi.org/10.1027/2151-2604/a000194

Green, C., & Coxhead, A. (2026). The Vocabulary Teachers' Competencies (VTC) framework: The core knowledge, skills and professional capabilities that support effective vocabulary instruction. *Literacy, 60*(1), e70016. https://doi.org/10.1111/lit.70016

Hsu, C.-C., & Sandford, B. A. (2007). The Delphi technique: Making sense of consensus. *Practical Assessment, Research, and Evaluation, 12*(10), 1–8. https://doi.org/10.7275/pdz9-th90

Ishikawa, A., Amagasa, M., Shiga, T., Tomizawa, G., Tatsuta, R., & Mieno, H. (1993). The max-min Delphi method and fuzzy Delphi method via fuzzy integration. *Fuzzy Sets and Systems, 55*(3), 241–253. https://doi.org/10.1016/0165-0114(93)90251-C

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement, 50*(1), 1–73. https://doi.org/10.1111/jedm.12000

Murray, T. J., Pipino, L. L., & van Gigch, J. P. (1985). A pilot study of fuzzy set modification of Delphi. *Human Systems Management, 5*(1), 76–80. https://doi.org/10.3233/HSM-1985-5111

Nation, I. S. P. (2022). *Learning vocabulary in another language* (3rd ed.). Cambridge University Press.

Opfer, V. D., & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of Educational Research, 81*(3), 376–407. https://doi.org/10.3102/0034654311413609

Richey, R. C., Fields, D. C., & Foxon, M. (2001). *Instructional design competencies: The standards* (3rd ed.). ERIC Clearinghouse on Information & Technology.

Schmidt, R. C. (1997). Managing Delphi surveys using nonparametric statistical techniques. *Decision Sciences, 28*(3), 763–774. https://doi.org/10.1111/j.1540-5915.1997.tb01330.x

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4–14. https://doi.org/10.3102/0013189X015002004

Teng, M. F., & Wong, W. L. (2026). Conceptualizing Vocabulary Teachers' Competencies (VTC) framework: A response to Green and Coxhead (2026). *Applied Language Sciences, 2*, 260112. https://doi.org/10.65553/ALS.260112

Wong, W. L., Teng, M. F., & Barrot, J. (2026). Developing and validating a vocabulary teaching competency questionnaire for teacher education in an Asia-focused context. *Innovation in Language Teaching and Learning*. https://doi.org/10.1080/17501229.2026.2751507

## Author bio

**Wei Lun Wong** is a senior lecturer at Faculty of Education, Universiti Kebangsaan Malaysia. His research interests include GenAI learner corpora, teacher education and applied linguistics. His works appear at *International Journal of Applied Linguistics, The Asia-Pacific Education Research, Research in Science and Technological Education, SAGE Open, Applied Linguistics, Innovation in Language Teaching and Learning* and *Scientific Reports* among others.
Email: colinw@ukm.edu.my
ORCID: https://orcid.org/0000-0002-4367-0157

**Lawrence Jun Zhang** is Professor of Education (Applied Linguistics/TESOL) and Associate Dean (AD) International for the Faculty of Arts and Education. His research programme spans cognitive, linguistic, sociocultural and developmental factors in reading/ biliteracy development, critical reading awareness in language education, metacognition, self-regulated learning (SRL) and reading development in L1 and L2 contexts, bilingual/biliteracy acquisition and bilingual/biliteracy education in primary and secondary schools, and learning and teaching English as a second/foreign language at university settings, the effects of self-regulated reading and writing instruction on bilingual/biliteracy development, emotions in learning and teaching, and teacher identity and cognition in language teacher education.
Email: lj.zhang@auckland.ac.nz
ORCID: https://orcid.org/0000-0003-1025-1746