International Journal of Multidisciplinary Research and Growth Evaluation.

Mental Toughness among Mandarin-Speaking Technology Innovators: Item Keying, Method Effects and a Qualitative Exploration of Intervention Experiences

Song-Ching Fan
Social and Public Administration School, Lingnan Normal University, Zhanjiang, Guangdong, China

* Corresponding Author: **Song-Ching Fan**

Article Info

ISSN (Online): 2582-7138
Volume: 07
Issue: 06
Received: 24-08-2026
Accepted: 26-09-2026
Published: 25-10-2026
Page No: 185-194

Abstract
Mental toughness is an individual-difference construct shaping how people appraise pressure. The 10-item Mental Toughness Questionnaire (MTQ10), developed in Western individualist settings, remains largely unexamined in Mandarin-speaking occupational samples. This exploratory mixed-methods study evaluated its measurement structure among Mandarin-speaking technology innovators and examined retrospective accounts of change following a positive psychological intervention. In September–October 2024, 216 adults—aspiring entrepreneurs (n = 87), technical professionals (n = 21), and innovation trainees (n = 108)—completed a Chinese-language MTQ10. Four competing models were compared using confirmatory factor analysis with full-information maximum likelihood estimation. The intended unidimensional model fitted poorly, $\chi^2(35) = 184.56$, CFI = .739, TLI = .665, RMSEA = .141, SRMR = .117. A two-factor model separating positively and negatively keyed items fitted substantially better, $\chi^2(34) = 61.24$, CFI = .953, TLI = .937, RMSEA = .061, SRMR = .057; $\Delta\chi^2(1) = 123.32$, $p < .001$. The factors correlated only .332, with separation reflecting item wording rather than content. In a substantive-factor model with an orthogonal keying factor, reverse-worded items loaded .173–.265 on the substantive factor but .463–.680 on the method factor. Composite reliability was .829 and .708 for positively and negatively keyed items, respectively. In July–August 2024, eight technology entrepreneurs who had completed two to seven PSI-based intervention sessions were interviewed. Grounded-theory analysis indicated movement from fear and self-doubt toward emotional regulation, self-efficacy, and proactive coping. In this sample, the MTQ10 primarily reflects positively worded items, making total scores difficult to justify. Findings are exploratory and hypothesis-generating.

DOI: https://doi.org/10.54660/.IJMRGE.2026.7.6.185-194

Keywords: Mental toughness, Cross-cultural measurement, Method effects, Item keying, Collectivism, Grounded theory

1. Introduction

Mental toughness describes a relatively stable individual difference in the capacity to remain purposeful, confident and emotionally steady when demands are high (Clough *et al*., 2008; Gucciardi *et al*., 2015) [5, 13]. Within the 4C framework — control, commitment, challenge and confidence — it predicts wellbeing, self-confidence and performance under pressure, and is generally treated as trainable rather than fixed (Dagnall *et al*., 2019; Zhu *et al*., 2025) [8, 35]. The 10-item Mental Toughness Questionnaire (MTQ10) is the most widely used short measure of the construct and has performed adequately across several cultural settings (Denovan *et al*., 2024; Farnsworth *et al*., 2022) [9, 11].

Like many brief self-report instruments, the MTQ10 mixes positively and negatively worded items. Where such instruments are administered outside the population in which they were developed, the mixture is a known source of difficulty: covariance that

is shared among items of the same keying, but not with the construct, can produce factors that look substantive and are not. The literature on the Rosenberg Self-Esteem Scale established both the phenomenon and the modelling strategy for detecting it (Marsh *et al*., 2010) [27], and method effects have already been reported for shortened toughness measures in cross-cultural work (Denovan *et al*., 2024) [9]. Whether the MTQ10 behaves this way in Mandarin-speaking samples has not been examined.

Technology innovators — founders, technical specialists and managers who convert technical expertise into new products and ventures — are an informative population for personality research on toughness. They work under high venture failure rates, chronic resource scarcity, frequent strategic reversals and public visibility of failure, while simultaneously carrying technical, leadership and investor-facing demands (Lima *et al*., 2020; Wang & Zhang, 2025; Yang, 2023) [22, 33, 34]. Psychological capital and resilience have been repeatedly linked to venture persistence in this group (Chen & Tao, 2021; Khooharo, 2022) [4, 18].

Whether a Western-derived toughness measure travels to this group in Chinese-speaking settings is not obvious. Resilience in Chinese cultural contexts is commonly enacted through role-based perseverance, endurance of hardship, family-centred support and concern for face rather than through assertions of personal control (Colonnello & Jannini, 2020; Liu, Huang, *et al*., 2025; Smith & Tian, 2014) [6, 23, 31]. Chinese founders are also more exposed to expectation and public evaluation than their Western counterparts (Song *et al*., 2025) [32]. Items that equate toughness with perceived personal agency may therefore under-represent the construct, and evidence on Chinese adaptations of resilience-related scales remains thin (Lei *et al*., 2025) [21]. Where an item additionally asks respondents to concede a shortfall — to endorse worry, hesitation or low enthusiasm — the normative cost of endorsement may be borne by all such items alike, which is precisely the condition under which keying-based method variance arises.

A second, related question concerns change. If mental toughness is malleable, structured psychological work should leave a describable trace in how people appraise pressure. The framework used here is the PSI model of positive psychological intervention, which specifies three components — P (positive reframing of beliefs), S (self-efficacy) and I (initiative and resistance to negative affect) — and was originally proposed as a psychological appraisal mechanism for mindset shift (Fan *et al*., 2021) [10]. To avoid confusion, PSI as used here denotes a positive psychological intervention framework and is distinct from Kuhl's Personality Systems Interactions theory, which shares the acronym but addresses volitional regulation of affect and cognition. The intervention examined here operationalises PSI through theta-based meditative practice: belief clearing, belief downloading and inner child work. Systematic review evidence indicates that meditative practice produces measurable neurobiological change in regions implicated in emotional regulation (Calderone *et al*., 2024) [3], and mindfulness-based programmes that combine meditation with cognitive restructuring have been shown to reduce stress and anxiety and to improve sleep quality and life satisfaction (Alvarado-Garcia *et al*., 2025) [1]. The practice examined here is conceptually consistent with these approaches, but it has not itself been evaluated in controlled trials, and it has not been evaluated in entrepreneurial samples at all. Resilience-oriented intervention research in Chinese-speaking populations has been dominated by mindfulness-based programmes and cognitive-behavioural approaches, the former targeting emotion regulation and stress tolerance and the latter cognitive restructuring and adaptive coping; the present work sits alongside these rather than in competition with them.

Self-report scales capture explicit, conscious appraisals and are therefore vulnerable to response bias (Krosnick *et al*., 2018; Mancuso *et al*., 2012) [20, 26]. Pairing them with in-depth interviews allows the meaning participants give to items and to their own change to be inspected directly (Liu, Fan, *et al*., 2025; Roulston, 2018) [24, 28]. We therefore combined psychometric evaluation with grounded-theory interviewing, integrating the two phases by conceptual triangulation: the interview material was used to interpret the quantitative result, and in particular to make sense of the structure recovered from the item set, rather than to corroborate it.

The study makes two contributions. It provides the first psychometric evaluation of the MTQ10 in a Mandarin-speaking innovation workforce, a population absent from the existing validation literature; and it pairs that evaluation with qualitative evidence about how respondents construe the construct, which allows the measurement finding to be read as a cultural response pattern rather than simply as psychometric noise.

Three questions guided the study. RQ1: What measurement structure does the MTQ10 show among Mandarin-speaking technology innovators, and is that structure attributable to item content or to item keying? RQ2: How do innovators who completed a PSI-based intervention retrospectively describe changes in mental toughness? RQ3: How do those who describe initially low toughness characterise their trajectory? Figure 1 sets out the conceptual framework. The design is exploratory: it cannot establish that the intervention caused the changes reported, and results are offered as hypothesis-generating.

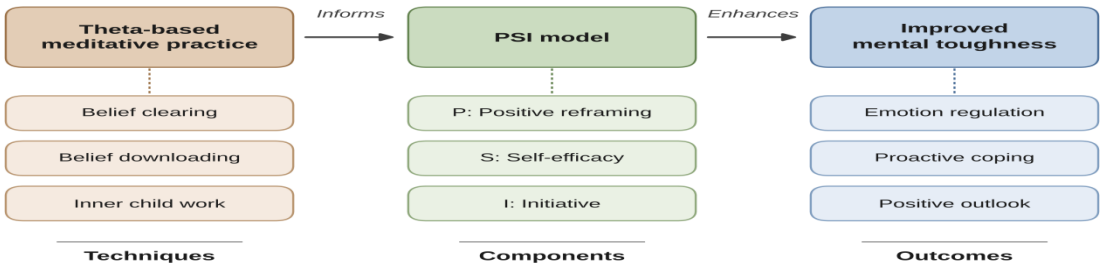


Fig 1: Conceptual Framework Linking Theta-Based Practice to the Three Psi Components and the Expected Mental Toughness Outcomes

2. Participants and Procedure

2.1. Design

An exploratory mixed-methods design was used, combining a cross-sectional psychometric phase with a retrospective qualitative phase. The quantitative phase evaluated the measurement structure and reliability of the MTQ10; the qualitative phase examined how entrepreneurs who had completed a PSI-based intervention described change in mental toughness. The two phases were conducted on separate samples and integrated at the interpretive level rather than statistically. Because intervention exposure was not recorded in the survey, the quantitative phase speaks only to measurement, and the interview material is used to interpret the measurement finding. No quantitative pre- and post-intervention data were collected, and the study therefore does not evaluate intervention outcomes.

2.2. Participants

Purposive sampling was used because technology innovators are heterogeneous and difficult to reach through probability frames. Quantitative participants were recruited from three innovation-related contexts representing successive points on the innovation continuum: aspiring entrepreneurs enrolled in government-sponsored entrepreneurship training in Taiwan (n = 87); technical professionals employed in Taiwanese technology firms (n = 21; the small n reflects restricted corporate access); and trainees of an innovation and entrepreneurship academy in Guangdong Province, China (n = 108).

The age item offered an “under 18” response band, which three respondents in the Guangdong trainee group selected. Because the survey was anonymous and recorded no identifying information, consent from a parent or legal guardian could neither be obtained nor verified for these respondents; their responses were therefore excluded from all analyses and from the deposited dataset. The analysed sample comprises 216 adults. One further respondent did not answer item A1; this case was retained and the single missing value was handled by full-information maximum likelihood. No other cases were excluded.

Sample size was determined by the number of eligible individuals accessible through the three participating settings during the data-collection window rather than by an a priori power analysis. The retained measurement model estimates 21 parameters in the covariance structure, giving a ratio of approximately 10 cases per estimated parameter, consistent with the minimum commonly applied to models of this size (Jobst *et al*., 2023; Kline, 2023) [17, 19]. For a model with 34 degrees of freedom at this sample size, the RMSEA test of close fit has power of approximately .67 to reject a true RMSEA of .08 and approximately .97 to reject a true RMSEA of .10 (MacCallum *et al*., 1996) [25]. The design is therefore well powered against gross misfit and only moderately powered against mild misfit, and the negative findings reported below should be read accordingly.

Of the 216 participants, 48.61% were men and 51.39% women; 72.69% lived in urban areas; 50.46% had a natural-science background and 37.50% a social-science background; and 81.94% had previously taken part in innovation or entrepreneurship activity. Full demographics are given in Table 1.

Table 1: Sample Characteristics of the Quantitative Phase

Variable	Category	N	%
Gender	Male	105	48.61
	Female	111	51.39
Age (years)	18–20	68	31.48
	21–23	33	15.28
	24 or above	115	53.24
Residence	Urban	157	72.69
	Rural	59	27.31
Field of study	Natural sciences	109	50.46
	Social sciences	81	37.50
	Arts and other	26	12.04
Only child	Yes	49	22.69
	No	167	77.31
Prior innovation or entrepreneurship activity	Yes	177	81.94
	No	39	18.06
Family background in technology entrepreneurship	Yes	58	26.85
	No	158	73.15

Note. N = 216. Percentages are column percentages within each variable and may not total 100 owing to rounding.

The three groups were pooled because the aim of this phase was measurement evaluation rather than group comparison, and because pooling theoretically related but heterogeneous respondents provides a stronger test of whether the latent structure holds across innovation-oriented settings. Two of the three groups were recruited in Taiwan (n = 108) and one in Guangdong Province, mainland China (n = 108). Pooling across these settings is defensible for the present purpose because participants completed the same Mandarin-language instrument and share the collectivist value orientations — role-based perseverance, familial obligation, concern for face — that motivate the cultural argument developed here. It is not, however, a substitute for evidence. Measurement invariance across the two settings was not tested, and formal multi-group comparison was not attempted because the unbalanced subgroup sizes, particularly the technical-professional subsample, would yield unstable parameter estimates. The analysis therefore establishes only what structure holds in the pooled sample, not that scores are directly comparable between settings. This constraint is revisited in the Discussion.

Qualitative participants (n = 8) were technology entrepreneurs who had completed two to seven sessions of a PSI-based mental toughness reshaping intervention and who

were still active in their ventures at interview. Sectors were communication technology (n = 2), environmental technology (n = 2), resource recycling (n = 1), online consulting (n = 2) and food technology (n = 1). Selection on intervention completion and continued venture activity introduces survivorship bias: these are persisters, not a representative cross-section of those who began.

2.3. Intervention

The intervention operationalised the three PSI components through theta-based meditative practice delivered in individual sessions. Each session opened with guided relaxation intended to induce a low-arousal, theta-dominant state, after which three techniques were applied: belief clearing (identifying and releasing limiting beliefs about failure, worth and capability), belief downloading (rehearsing alternative appraisals aligned with positive reframing, self-efficacy and initiative), and inner child work (reprocessing early emotional experiences implicated in present avoidance). Sessions were scheduled at participants' request rather than on a fixed protocol, which accounts for the range of two to seven sessions. The intervention was not manualised and no fidelity checks were applied; this limits replicability and is revisited in the Discussion.

2.4. Measures

The MTQ10 (Dagnall *et al*., 2019) [8] comprises 10 items, six positively and four negatively worded, rated on a seven-point Likert scale; negatively keyed items were reverse-scored before analysis so that higher scores indicate greater toughness throughout. The Chinese-language version administered here was prepared by the author, who translated the published English items and adapted the wording where a literal rendering would not carry the intended sense for Chinese-speaking respondents. No independent forward and back translation, expert panel review or cognitive pretesting was undertaken, and the translation had not been validated before this study. The instrument evaluated here is therefore an author-adapted Chinese version of the MTQ10 rather than an independently validated translation, and the structural findings reported below apply to that version. This is revisited in the Discussion. The questionnaire was self-administered and anonymous, and no identifying data were recorded. Item descriptors, descriptive statistics and standardised loadings are reported in Table 2.

The interview schedule was semi-structured and covered three domains: functioning before the intervention, experience of the intervention itself, and current appraisal of pressure and setbacks.

2.5. Procedure

Data were collected in two phases. The survey ran from September to October 2024 and the interviews from 12 July to 23 August 2024, so interviewees were approximately one to three years post-intervention at the time of interview. The two phases were independent and were not sequenced; the ordering of their labels reflects the structure of the report rather than the order of data collection. All materials were administered in Mandarin. Interviews lasted approximately 30–40 minutes, were audio-recorded with consent and were transcribed verbatim.

2.6. Analysis

Confirmatory factor analysis was conducted using full-information maximum likelihood estimation on the 10 MTQ10 items, with means and intercepts estimated; one item response was missing and no cases were excluded on that account. Rather than testing a single a priori structure, four specifications suggested by the measurement literature on brief self-report scales that mix item keying were compared. Model 1 specified the single common factor implied by the original scoring of the MTQ10. Model 2 specified two correlated factors corresponding to the positively and the negatively keyed items. Model 3 specified a single substantive factor across all 10 items together with an orthogonal method factor loading only on the four negatively keyed items, which represents item keying as a source of common variance that is not part of the construct (Denovan *et al*., 2024; Marsh *et al*., 2010) [9, 27]. Model 4 replaced the method factor with correlated residuals among the four negatively keyed items, a more conservative representation of the same phenomenon.

Model fit was judged by the model chi-square, the chi-square to degrees-of-freedom ratio, CFI, TLI, RMSEA with its 90% confidence interval and the test of close fit, and SRMR; models were compared on AIC, and Model 1 was additionally compared with Model 2 by likelihood ratio test, the two being nested because Model 1 is Model 2 with the interfactor correlation fixed at unity. Composite reliability (CR) and average variance extracted (AVE) were computed from the standardised loadings of the retained model and interpreted against the criteria of Fornell and Larcker (1981) [12] and Hair *et al*. (2011) [14]; internal consistency was indexed by Cronbach's α and McDonald's ω. Discriminant validity was assessed by comparing the square root of AVE with the interfactor correlation. Confidence intervals for standardised loadings were obtained from 1,000 bias-corrected bootstrap samples drawn from the complete cases. All analyses were conducted in Python 3 using the semopy package.

Interview transcripts were analysed in NVivo 14 following grounded-theory procedures — open, axial and selective coding. The qualitative phase was designed to explore experiential and process-level accounts of how mental toughness is construed and reported to change, not to validate or confirm the PSI model; its conclusions therefore concern process-level coherence rather than model accuracy. Coding decisions and word frequencies were documented to support transparency (Saldaña, 2021; Schreier, 2012) [29, 30]. Coding proceeded until theme redundancy and conceptual completeness were reached; given the homogeneity of the participants and the focused research questions, saturation was judged to have been achieved within the eight cases. The full coding structure is provided as supplementary material.

Table 2: Item Descriptives and Standardised Loadings for the Retained Two-Factor Model

Item	Key	Content (abbreviated)	M	SD	λ	95% CI	R^2
A4	+	Coping with problems that emerge	4.83	1.08	.758	[.627, .850]	.574
A5	+	Perceived personal worth	5.30	1.14	.718	[.569, .803]	.516
A8	+	Sense of being in control	4.74	1.11	.692	[.565, .786]	.479
A1	+	Remaining calm under significant pressure	5.08	1.10	.655	[.526, .764]	.429
A10	+	Perceiving the brighter side of life	5.24	1.14	.638	[.499, .736]	.407
A9	+	Responding quickly to unexpected situations	5.27	1.24	.537	[.334, .663]	.289
A2	−	Worrying about things in advance	4.38	1.48	.740	[.601, .874]	.548
A6	−	Feeling at a loss with competing demands	4.44	1.52	.645	[.497, .768]	.416
A7	−	Prolonged worry after making a mistake	4.14	1.57	.561	[.392, .694]	.314
A3	−	Difficulty feeling enthusiastic about obligatory tasks	4.47	1.36	.501	[.316, .646]	.251

Note. Key indicates positively (+) or negatively (−) keyed items; negatively keyed items were reverse-scored before analysis, so a higher mean indicates greater toughness for every item. Content is an abbreviated descriptor, not the item as administered. Means are on the seven-point response scale. For A1, n = 215; for all other items, n = 216. λ = standardised factor loading; all loadings $p < .001$. Confidence intervals are bias-corrected bootstrap intervals based on 1,000 resamples of the complete cases. Positively keyed items: α = .820, ω = .829, CR = .829, AVE = .449. Negatively keyed items: α = .703, ω = .708, CR = .708, AVE = .382. Interfactor correlation r = .332. An AVE between .40 and .50 is regarded as tolerable where CR exceeds .70 (Hair *et al*., 2011) [14].

3. Results

Structural evaluation (RQ1): Fit statistics for the four competing models are reported in Table 3. The single-factor model implied by the original scoring of the MTQ10 fitted poorly on every index, $\chi^2(35, N = 216) = 184.56$, $p < .001$, $\chi^2/df = 5.27$, CFI = .739, TLI = .665, RMSEA = .141, 90% CI [.121, .161], SRMR = .117, AIC = 244.56. The instrument is not unidimensional in this sample.

Model 2, which distinguished the positively from the negatively keyed items, fitted acceptably, $\chi^2(34, N = 216) = 61.24$, $p = .003$, $\chi^2/df = 1.80$, CFI = .953, TLI = .937, RMSEA = .061, 90% CI [.035, .085], PCLOSE = .216, SRMR = .057, AIC = 123.24. Because Model 1 is Model 2 with the interfactor correlation fixed at unity, the two are nested, and the difference was decisive, $\Delta\chi^2(1) = 123.32$, $p < .001$. All 10 standardised loadings were significant and ranged from .501 to .758, and squared multiple correlations from .251 to .574 (Table 2).

Table 3: Fit Statistics for Four Competing Measurement Models of the MTQ10

Model	χ^2	df	p	χ^2/df	CFI	TLI	RMSEA [90% CI]	SRMR	AIC
1. Single factor	184.56	35	< .001	5.27	.739	.665	.141 [.121, .161]	.117	244.56
2. Two factors by item keying	61.24	34	.003	1.80	.953	.937	.061 [.035, .085]	.057	123.24
3. Substantive factor + keying method factor	60.33	31	.001	1.95	.949	.926	.066 [.041, .091]	.056	128.33
4. Single factor + correlated residuals	59.60	29	< .001	2.06	.947	.917	.070 [.044, .095]	.056	131.60

Note. N = 216. Baseline (independence) model $\chi^2(45) = 618.85$, $p < .001$. All models estimated by full-information maximum likelihood with means and intercepts free; one item response was missing and no cases were excluded on that account. Model 1 is nested within Model 2, the former being the latter with the interfactor correlation fixed at unity, $\Delta\chi^2(1) = 123.32$, $p < .001$. For Model 2, PCLOSE = .216 and Hoelter's critical N at .05 was 171. Lower AIC indicates better fit relative to the other models in the set.

What separates the two factors is not item content. The partition is exact: all six positively keyed items loaded on the first factor and all four negatively keyed items on the second, with no item crossing the boundary. The two factors correlated .332, far from the unity that a single dimension would require and low for two facets of one disposition. Discriminant validity was accordingly satisfied, the square roots of AVE (.670 and .618) both exceeding the interfactor correlation, but it is satisfied by a distinction that has no substantive content.

Model 3 tested this reading by specifying one substantive factor across all 10 items together with an orthogonal method factor loading only on the negatively keyed items. It fitted comparably, $\chi^2(31, N = 216) = 60.33$, $p = .001$, CFI = .949, TLI = .926, RMSEA = .066, 90% CI [.041, .091], SRMR = .056, though at a cost in parsimony that left it behind on AIC (128.33). Model 4, which replaced the method factor with correlated residuals among the reverse-worded items, was similar, $\chi^2(29, N = 216) = 59.60$, $p < .001$, CFI = .947, TLI = .917, RMSEA = .070, 90% CI [.044, .095], SRMR = .056, AIC = 131.60. On fit alone, Model 2 is retained.

The parameter estimates of Model 3 are nonetheless informative, and all were admissible, with no negative residual variances. In that solution the six positively keyed items loaded .537 to .761 on the substantive factor (A4 = .761, A5 = .715, A8 = .693, A1 = .658, A10 = .635, A9 = .537), whereas the four negatively keyed items loaded only .173 to .265 on it (A2 = .265, A7 = .186, A3 = .184, A6 = .173) while loading .463 to .680 on the method factor (A2 = .680, A6 = .638, A7 = .531, A3 = .463). Between 3% and 7% of the variance in the reverse-worded items is attributable to the substantive construct. Taken with the exactness of the keying partition and the low interfactor correlation in Model 2, this indicates that the second factor in Model 2 carries variance generated by item wording rather than by mental toughness, and that in this sample the MTQ10 measures the construct almost entirely through its positively worded items.

Reliability followed the same division. For the six positively keyed items, α = .820, ω = .829, CR = .829 and AVE = .449; for the four negatively keyed items, α = .703, ω = .708, CR = .708 and AVE = .382. Both AVE values fall below the conventional .50 threshold; values between .40 and .50 are regarded as tolerable where CR exceeds .70 (Fornell & Larcker, 1981; Hair *et al*., 2011) [12, 14], but the negatively keyed set falls below even that. Item A3, which concerns difficulty feeling enthusiastic about obligatory tasks, was the weakest indicator in the instrument, λ = .501, R^2 = .251.

One further index qualifies the result. For the retained model, Hoelter's critical N at the .05 level was 171, below both the achieved sample size and the conventional benchmark of 200, indicating that the significant model chi-square is not attributable to sample size alone.

Qualitative findings (RQ2 and RQ3): Grounded-theory analysis produced five selective categories, summarised below and set out in full in the supplementary material. The findings are presented as illustrative process patterns within specific innovation contexts rather than as evidence of intervention effectiveness or generalisable outcomes.

Low mental toughness before intervention: All eight participants described a period of low toughness marked by fear of failure, self-doubt and anxiety, with corresponding reluctance to take risks. Family background featured prominently: dependence on others, family expectation and early experience were spontaneously raised as sources of constraint. Five participants located themselves in this state at the point they entered the intervention.

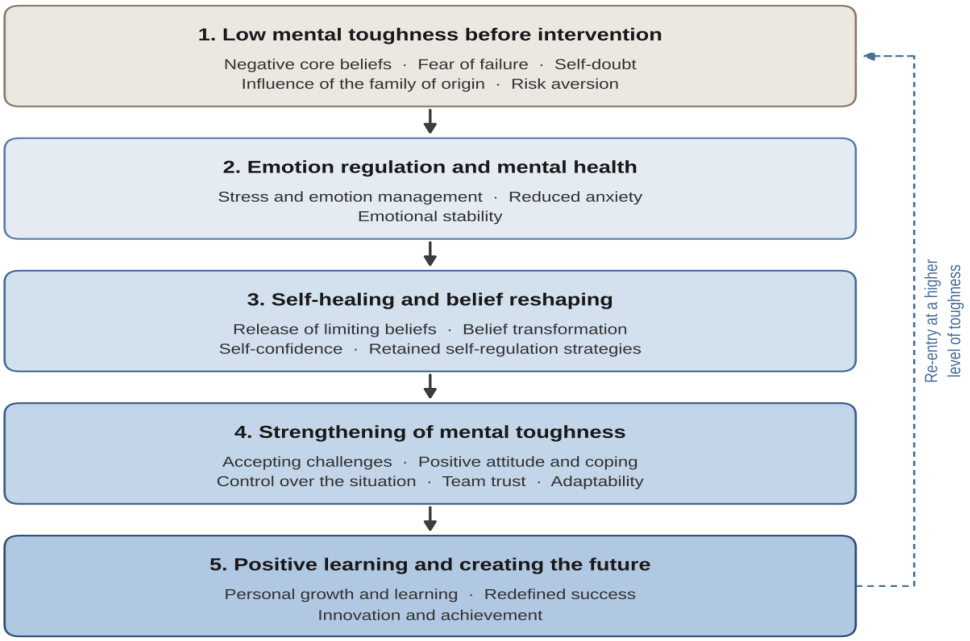
Emotion regulation and mental health: Stress was the most frequent code in the corpus (28 mentions). Before the intervention, participants described stress as impairing problem-solving and daily functioning. Afterwards they reported greater self-awareness, calmness under pressure and reappraisal of demands as manageable — a pattern consistent with evidence that emotion-regulation capacity supports resilience in high-pressure technical work (Hopman *et al.*, 2023; Jentsch & Wolf, 2024) [15, 16].

Self-healing and belief reshaping: Participants attributed change to the release of limiting beliefs and the installation of alternatives. Self-confidence was the second most frequent code (18 mentions). Six participants reported adopting concrete strategies — meditation, deep breathing, cognitive reframing, structured problem-solving — that they had retained. All eight described reduced fear of failure.

Strengthening of mental toughness: Seven of eight reported greater willingness to accept challenges (accepting challenges, 11 mentions) and improved handling of unexpected events, shifting from avoidance to proactive engagement. Participants also emphasised situational control, team trust and acceptance of difference as resources they now drew on deliberately.

Positive learning and creating the future: Learning and growth (15 mentions) dominated accounts of the present, with participants emphasising self-directed development and redefined criteria for success rather than past failure. Innovation, venture creation and reframed definitions of success recurred across transcripts. One participant summarised the shift as treating failure as part of growth rather than as evidence of inadequacy.

Across accounts, stress was attributed primarily to external collective obligations — family expectation, social evaluation — rather than to deficits of personal control. Figure 2 organises these categories into a provisional transformation model.



Note. Arrows denote hypothesised rather than demonstrated sequences. The model is derived from eight retrospective accounts and has not been tested prospectively.

Fig 2: Provisional Transformation Model of Mental Toughness Development Derived from the Grounded-Theory Analysis

4. Discussion

4.1. Measurement Implications

The answer to RQ1 is qualified in a specific and informative way. The MTQ10 did not behave as a unidimensional instrument in this sample, and the unidimensional model its scoring assumes was decisively rejected, $\Delta\chi^2(1) = 123.32$, $p < .001$. But the departure from unidimensionality did not follow the content of the items. It followed their keying. Every positively worded item grouped together and every negatively worded item grouped together, the resulting factors correlated only .332, and in a model that represents keying explicitly the reverse-worded items retained loadings of only .173 to .265 on the substantive factor while loading .463 to .680 on the method factor. Method effects of exactly this kind have been documented for instruments that mix item keying (Marsh *et al.*, 2010) [27] and reported for shortened

toughness measures in cross-cultural work (Denovan *et al*., 2024) [9]; the present data extend that observation to a Mandarin-speaking occupational sample.

The interview material offers a reading of why the effect might be pronounced here. Across the eight accounts, difficulty was attributed to collective obligation — family expectation, social evaluation, responsibility to others — rather than to any deficit of personal control. In a setting where perseverance is enacted as role fulfilment, the negatively keyed items do not function as neutral self-descriptions. To endorse difficulty in feeling enthusiastic about tasks one is obliged to complete is not to report a low disposition; it is to concede a failure of obligation. The four negatively keyed items ask respondents to make such a concession in four different ways, and it is plausible that a shared reluctance to do so produces the common variance that the method factor captures. Consistent with this, the weakest indicator in the whole instrument, item A3, is the item that states the concession most directly, $\lambda = .501$, $R^2 = .251$, falling to .184 on the substantive factor in Model 3.

This is presented as an interpretation, not as a demonstration. No cognitive interviewing was conducted on the items themselves, the qualitative and quantitative phases were drawn from separate samples, and a method factor is a statistical description compatible with several explanations — acquiescence, careless responding to reversals, and socially normative concession avoidance among them. A further possibility is properly linguistic rather than psychological: the Chinese version was prepared and adapted by the author without independent back-translation, so systematic differences in how the four negatively worded items were rendered could contribute to the shared variance the method factor captures. Distinguishing these requires evidence the present design cannot supply, and independent translation review is a precondition for doing so.

Two practical consequences follow nonetheless. First, a simple MTQ10 total score is not defensible in samples of this kind. Summing across items that load on different factors, and doing so when those factors correlate at .332 and four of the ten items carry almost no construct variance, produces a composite whose meaning varies with the balance of keying rather than with the trait. Users should model the keying effect explicitly, or score the positively and negatively keyed items separately, until invariance of the full instrument has been established. Second, the convergent validity shortfall is concentrated where the keying effect is: AVE was .449 among the positively keyed items and only .382 among the negatively keyed items. Remediation should target the reverse-worded items specifically, whether by re-translation, by replacement with positively worded equivalents, or by retaining them with an explicit method factor in the measurement model.

Establishing that a toughness measure is culturally adequate therefore requires more than acceptable global fit — a point these results make concretely, since the two-factor model fitted well by every conventional index while resting on a distinction that has no substantive content. Cognitive interviewing to surface indigenous expressions of the construct, item response theory and differential item functioning analysis to test whether items behave equivalently at matched trait levels, and configural, metric and scalar invariance testing across contrasting Mandarin-speaking settings are all needed before cross-cultural score comparisons can be trusted (Denovan *et al*., 2024; Lei *et al*., 2025) [9, 21]. Beyond testing existing items, culturally responsive scale development might integrate constructs such as role-based perseverance, relational responsibility and socially embedded self-regulation alongside the conventional dimensions of control and commitment, and might do so without relying on reverse wording to achieve balance.

4.2. Interpreting the Reported Change

The qualitative findings converge on emotional regulation, belief restructuring and self-efficacy as the descriptive core of reported change, which is consistent with social cognitive accounts of agency (Bandura, 2001) [2] and with intervention literature on positive reframing. The changes were still being described one to three years later, which is worth noting given how often intervention effects are measured only at termination. But the design cannot license a causal reading. Participants were selected on completion and continued venture activity; no baseline was taken; there was no comparison condition; and the accounts are retrospective and therefore open to reconstruction. Maturation, concurrent entrepreneurial learning, expectancy and allegiance effects are all live alternatives, and the last of these is reinforced by the fact that the interviews were conducted by a researcher connected to the intervention. Effect sizes routinely shrink from uncontrolled pilots to randomised trials (Cristea *et al*., 2017) [7], and there is no reason to expect an exception here.

4.3. Limitations

Six limitations bear on interpretation. First, intervention exposure was not measured in the survey, so the two phases cannot be linked statistically; the quantitative phase speaks only to measurement. Second, the instrument did not function unidimensionally in this sample and the departure followed item keying; results are therefore reported at subscale level rather than as a composite, and the study provides no basis for comparing MTQ10 total scores between this population and the Western samples in which the instrument was developed. Relatedly, half the sample was recruited in Taiwan and half in Guangdong, and measurement invariance across the two settings was not tested; the pooled solution does not license comparison of scores between them, and pooling may obscure differences between subgroups at different entrepreneurial stages or in different professional roles. Third, the Chinese version of the instrument was translated and adapted by the author without independent back-translation, expert review or cognitive pretesting; the findings therefore describe that version rather than the MTQ10 as published, and translation effects cannot be separated from the keying effect reported here. Fourth, the four measurement models were compared in a single sample without cross-validation, and the superiority of the retained model therefore requires replication before it is treated as established; the design is also only moderately powered against mild misfit. Fifth, the cross-sectional design gives no information about the temporal stability of the construct. Sixth, the author is professionally connected to the organisation that delivered the intervention and conducted the interviews personally, which may have encouraged favourable retrospective accounts; this is declared below and should be weighed when reading the qualitative findings.

4.4. Implications and Future Directions

Within these limits the findings suggest testable propositions. Interventions targeting belief restructuring and emotion

regulation may be of greatest benefit to innovators who enter with low toughness, and culturally adapted programmes may need to address collectivist stressors — family obligation, role pressure, public evaluation — rather than generic stress alone. Both propositions require randomised or propensity-matched designs with pre-intervention baselines, active comparison conditions such as stress education or mindfulness, prospective follow-up, independent interviewers and systematic tracking of dropout. On the measurement side, the most immediate next step is to replicate the model comparison reported here in an independent Mandarin-speaking sample and to test configural, metric and scalar invariance of the MTQ10 across Mandarin-speaking societies with the keying effect modelled explicitly. Until such work exists, resilience programming in innovation training and technology organisations should be treated as experimental and evaluated accordingly.

5. Conclusions

Among Mandarin-speaking technology innovators the MTQ10 does not function as the unidimensional instrument its scoring assumes. The items separate along keying rather than content, the resulting factors correlate only weakly, convergent validity is below convention and weakest among the reverse-worded items, and in a model that represents keying explicitly those items carry almost no construct variance — a pattern better described as a method effect than as evidence of two components of mental toughness. Total scores should therefore not be used in this population, and the instrument should not be used for cross-cultural comparison until the effect has been modelled and invariance established. Innovators who had completed a PSI-based intervention consistently described a movement from fear of failure and self-doubt towards emotional regulation, self-efficacy and proactive coping, sustained one to three years later; the design does not establish that the intervention produced this. Taken together, the two phases indicate that mental toughness in this population is better understood as a culturally embedded form of adaptive functioning than as a context-free trait, and they define the measurement and trial work that would be needed to test that claim properly.

Declarations

Ethical Considerations

The study was reviewed and approved by the Ethics Committee of Lingnan Normal University, Zhanjiang, Guangdong, China (approval no. LNU-2026-S0303), with approval granted on 18 March 2026, and was conducted in accordance with the Declaration of Helsinki. Two matters are stated explicitly. First, the timing of the review: data were collected between July and October 2024, and the committee's review was obtained retrospectively, after data collection was complete, when the study was being prepared for publication. At the time of data collection each participant received written information on the purpose of the study, anonymity, confidentiality and the right to withdraw at any time without consequence, and each gave written informed consent before taking part; no identifiable personal data were collected in the survey phase. The committee reviewed the completed study, including the participant information and consent materials used at the time, and confirmed that it satisfied the institution's ethical requirements and the requirements of the Measures for Ethical Review of Life Science and Medical Research Involving Humans (2023). Second, the exclusion of three respondents: the survey's age item offered an “under 18” response band, which three respondents selected. Because the survey was anonymous and recorded no identifying information, consent from a parent or legal guardian could neither be obtained nor verified for these respondents. Their responses were therefore excluded from all analyses and from the deposited dataset, and the analysed sample comprises 216 adult participants, each of whom gave written informed consent. This exclusion was reported in writing to the Ethics Committee of Lingnan Normal University. Copies of the approval certificate, the committee's acknowledgement of the exclusion and the signed consent forms can be supplied to the editors on request. The study is not a clinical trial under the ICMJE definition: participants were not prospectively assigned to an intervention by the investigator and no health outcome was measured, so trial registration does not apply.

Consent to Participate and For Publication

Written informed consent to participate was obtained from all participants in both the survey and the interview phases before data collection. All interview participants gave written consent for anonymised material derived from their accounts to be published. No individually identifiable data are reported in this article.

Declaration of Conflicting Interests

The author is professionally connected to the organisation that delivered the PSI-based intervention examined in the qualitative phase of this study, and personally conducted the eight interviews. This constitutes a non-financial competing interest. It is disclosed here, is noted in the Limitations section, and should be weighed by readers when interpreting the retrospective interview accounts. The author declares no financial conflicts of interest with respect to the research, authorship or publication of this article.

Funding

This work was supported by the Guangdong Provincial Department of Science and Technology (grant no. 2024A0505050017). The funder had no role in the design of the study, in the collection, analysis or interpretation of data, in the writing of the manuscript, or in the decision to submit it for publication.

Data Availability

The de-identified quantitative dataset (216 cases, with the three excluded respondents removed), the item-level codebook and the confirmatory factor analysis syntax for all four measurement models are available from the author on reasonable request. The interview transcripts contain contextual detail through which the eight participants could be identified, and participants did not consent to public release of their accounts; the transcripts are therefore held under protected access. Qualified researchers may request access by contacting the author with a research plan and a signed confidentiality undertaking, and requests will be assessed by the author's institutional ethics committee. The full grounded-theory coding structure derived from the transcripts is provided as supplementary material.

Preregistration

This study was not preregistered.

Acknowledgments

The author thanks the participating institutions and all respondents and interviewees for their time. Data were collected and interviews conducted in Mandarin; as the author is not a native speaker of English, AI-assisted language tools were used for translation from Mandarin and for language editing of the manuscript. No AI tool was used for study design, data analysis, data interpretation or the generation of content. The author reviewed and edited all output and takes full responsibility for the content of the article.

References

1. Alvarado-Garcia PAA, Soto-Vasquez MR, Infantes Gomez FM, Guzman Rodriguez NM, Castro-Paniagua WG. Effect of a mindfulness program on stress, anxiety, depression, sleep quality, social support, and life satisfaction: a quasi-experimental study in college students. Front Psychol 2025;16:1508934. https://doi.org/10.3389/fpsyg.2025.1508934
2. Bandura A. Social cognitive theory: an agentic perspective. Annu Rev Psychol 2001;52(1):1-26. https://doi.org/10.1146/annurev.psych.52.1.1
3. Calderone A, Latella D, Impellizzeri F, De Pasquale P, Fama F, Quartarone A. Neurobiological changes induced by mindfulness and meditation: a systematic review. Biomedicines 2024;12(11):2613. https://doi.org/10.3390/biomedicines12112613
4. Chen H, Tao Y. Efficacy of entrepreneurs' psychological capital on the performance of new ventures in the development of regional economy in the Greater Bay Area. Front Psychol 2021;12:705095. https://doi.org/10.3389/fpsyg.2021.705095
5. Clough P, Earle K, Strycharczyk D. Developing resilience through coaching: MTQ48. In: Passmore J, editor. Psychometrics in coaching: using psychological and psychometric tools for development. London: Kogan Page; 2008. p. 208-23.
6. Colonnello E, Jannini EA. Impact of Chinese traditional culture and related social norms on current Chinese sexuality. In: Cultural differences and the practice of sexual medicine. Springer; 2020. p. 95-113. https://doi.org/10.1007/978-3-030-36222-5_6
7. Cristea IA, Gentili C, Cotet CD, Palomba D, Barbui C, Cuijpers P. Efficacy of psychotherapies for borderline personality disorder: a systematic review and meta-analysis. JAMA Psychiatry 2017;74(4):319-328. https://doi.org/10.1001/jamapsychiatry.2016.4287
8. Dagnall N, Denovan A, Papageorgiou KA, Clough PJ, Parker A, Drinkwater KG. Psychometric assessment of shortened mental toughness questionnaires (MTQ): factor structure of the MTQ-18 and the MTQ-10. Front Psychol 2019;10:1933. https://doi.org/10.3389/fpsyg.2019.01933
9. Denovan A, Dagnall N, Artamonova E, Papageorgiou KA. Evaluation of the 10-item mental toughness questionnaire (MTQ10): cross-cultural assessment and scrutiny of method effects. Curr Psychol 2024;43:1-13. https://doi.org/10.1007/s12144-024-06562-9
10. Fan SC, Shih HC, Tseng HC, Chang KF, Li WC, Shia AS. Self-efficacy triggers psychological appraisal mechanism for mindset shift. Int J Ment Health Promot 2021;23:57-73. https://doi.org/10.32604/ijmhp.2021.012177
11. Farnsworth JL, Marshal A, Myers NL. Mental toughness measures: a systematic review of measurement properties for practitioners. J Appl Sport Psychol 2022;34(3):479-494. https://doi.org/10.1080/10413200.2020.1866710
12. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. J Mark Res 1981;18(1):39-50. https://doi.org/10.1177/002224378101800104
13. Gucciardi DF, Hanton S, Gordon S, Mallett CJ, Temby P. The concept of mental toughness: tests of dimensionality, nomological network, and traitness. J Pers 2015;83(1):26-44. https://doi.org/10.1111/jopy.12079
14. Hair JF, Ringle CM, Sarstedt M. PLS-SEM: indeed a silver bullet. J Mark Theory Pract 2011;19(2):139-152. https://doi.org/10.2753/MTP1069-6679190202
15. Hopman K, Richards D, Norberg MM. A digital coach to promote emotion regulation skills. Multimodal Technol Interact 2023;7(6):57. https://doi.org/10.3390/mti7060057
16. Jentsch VL, Wolf OT. The bidirectional interplay between stress and cognitive emotion regulation. Psychoneuroendocrinology 2024;160:106719. https://doi.org/10.1016/j.psyneuen.2023.106719
17. Jobst LJ, Bader M, Moshagen M. A tutorial on assessing statistical power and determining sample size for structural equation models. Psychol Methods 2023;28(1):207-226. https://doi.org/10.1037/met0000327
18. Khooharo TA. Role of psychological capital in dealing with the business challenges faced by entrepreneurs. J Contemp Issues Bus Gov 2022;28(3). https://doi.org/10.47750/cibg.2022.28.03.017
19. Kline RB. Principles and practice of structural equation modeling. 5th ed. Guilford Press; 2023.
20. Krosnick JA, Judd CM, Wittenbrink B. The measurement of attitudes. In: The handbook of attitudes: Vol. 1. Basic principles. Routledge; 2018. p. 45-105.
21. Lei Y, Duan C, Shen K. Development and validation of the Chinese Mental Health Value Scale: a tool for culturally-informed psychological assessment. Front Psychol 2025;16:1417443. https://doi.org/10.3389/fpsyg.2025.1417443
22. Lima LGD, Nassif VMJ, Garcon MM. The power of psychological capital: the strength of beliefs in entrepreneurial behavior. Rev Adm Contemp 2020;24:317-334. https://doi.org/10.1590/1982-7849rac2020180226
23. Liu J, Huang Q, Park HJ, Qin DB. Beyond collectivism and individualism: Chinese international students' perceptions of contemporary child-centered parenting. Fam Relat 2025. https://doi.org/10.1111/fare.13157
24. Liu TF, Fan SC, Jiang XL. How can qualitative in-depth interviews optimize cross-cultural measurement of academic resilience? Front Psychol 2025;16:1444978. https://doi.org/10.3389/fpsyg.2025.1444978
25. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. Psychol Methods 1996;1(2):130-149. https://doi.org/10.1037/1082-989X.1.2.130
26. Mancuso CA, Ranawat AS, Meftah M, Koob TW, Ranawat CS. Properties of the patient administered

questionnaires. J Arthroplasty 2012;27(4):575-582. https://doi.org/10.1016/j.arth.2011.07.014
27. Marsh HW, Scalas LF, Nagengast B. Longitudinal tests of competing factor structures for the Rosenberg Self-Esteem Scale: traits, ephemeral artifacts, and stable response styles. Psychol Assess 2010;22(2):366-381. https://doi.org/10.1037/a0019225
28. Roulston K. Qualitative interviewing and epistemics. Qual Res 2018;18(3):322-341. https://doi.org/10.1177/1468794117721738
29. Saldaña J. The coding manual for qualitative researchers. 4th ed. Sage; 2021.
30. Schreier M. Qualitative content analysis in practice. Sage; 2012.
31. Smith WK, Tian Y. Individual motivations to create social enterprises: cultural impacts in China. Acad Manag Proc 2014;2014:13205. https://doi.org/10.5465/ambpp.2014.13205abstract
32. Song C, Nahm AY, Song Z. Childhood socioeconomic status of Chinese entrepreneurs and company risk-taking. Int Rev Econ Finance 2025;98:103958. https://doi.org/10.1016/j.iref.2025.103958
33. Wang R, Zhang G. Does optimising the business environment enhance enterprise resilience? The role of digital innovation and entrepreneurial spirit. Humanit Soc Sci Commun 2025;12(1):1-15. https://doi.org/10.1057/s41599-025-04704-3
34. Yang A. Psychological resilience of entrepreneurial failure: an application of positive psychology to entrepreneurial failure repair. Acad J Manag Soc Sci 2023;5(2):17-21. https://doi.org/10.54097/ajmss.v5i2.06
35. Zhu C, Peng L, Li S. The impact of mental toughness on subjective well-being among Chinese college students in the post-COVID-19 era. Front Public Health 2025;13:1556096. https://doi.org/10.3389/fpubh.2025.1556096

How to Cite This Article
Fan SC. Mental toughness among Mandarin-Speaking technology innovators: item keying, method effects and a qualitative exploration of intervention experiences. Int J Multidiscip Res Growth Eval. 2026;7(6):185-194. doi:10.54660/.IJMRGE.2026.7.6.185-194.

Creative Commons (CC) License
This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms