

A Reading of the Behavioral Voice Motivators of Internal Auditors in Developing Effective Voice

Seyed Mehdi. Teimorian¹, Mehdi. Khalilpour^{2*}, Javad. Ramezani³, Mehdi. Safari Gerayli⁴

¹ Department of Accounting, No.c, Islamic Azad University, Nour, Iran

² Assistant Professor, Department of Accounting, No.c, Islamic Azad University, Nour, Iran

³ Associate Professor, Department of Accounting, No.c, Islamic Azad University, Nour, Iran

⁴ Associate Professor, Department of Accounting, BG.C, Islamic Azad University, Bandargaz, Iran

* Corresponding author email address: khalilpournmehdi@iau.ac.ir

E d i t o r	R e v i e w e r s
Azar Kafashpoor Professor, Department of Educational Management and Human Resource Development, Ferdowsi University of Mashhad, Mashhad, Iran kafashpor@um.ac.ir	Reviewer 1: Abbas Monavarian Professor, Management Department, Tehran University, Tehran, Iran. Email: amonavar@ut.ac.ir Reviewer 2: Rezvan Hosseingholizadeh Associate Professor, Department of Educational Management and Human Resource Development, Ferdowsi University of Mashhad, Mashhad, Iran. Email: rhgholizadeh@um.ac.ir

1. Round 1

1.1. Reviewer 1

Reviewer:

The literature review would benefit from a sharper articulation of the research gap. The introduction currently states that “relatively little empirical attention has been devoted to identifying how internal auditors’ own psychological and behavioral voice motivators translate into greater internal audit effectiveness.” This gap should be demonstrated more systematically rather than asserted. The authors should distinguish what is already known from Ebrahimi et al. and other internal-audit studies from what is genuinely novel in the present study. In particular, because Ebrahimi et al. examined “external and internal drivers of individual behavioral voice” in relation to moral courage and internal auditor effectiveness, the manuscript must explain precisely how the present model differs conceptually, empirically, contextually, and methodologically from that prior work.

The manuscript should improve the alignment between its stated novelty and the actual design. The abstract claims that “this is the first study that specifically focuses on evaluating the effect of behavioral voice motivators on internal audit effectiveness.” Such a strong priority claim requires comprehensive evidence and should generally be avoided unless supported by a systematic

literature search. The authors should replace this with a more defensible statement, such as indicating that the study contributes to a relatively underexplored stream of research linking individual psychological resources associated with voice to internal audit effectiveness, particularly in the Tehran Stock Exchange context.

The definition of the study population is currently imprecise. The methods section states that “the statistical population consisted of internal auditors working in companies listed on the Tehran Stock Exchange,” whereas the Persian text refers to “internal auditors working in the Tehran Stock Exchange,” which could be interpreted as employees of the Exchange itself. The authors should clearly specify whether participants were internal auditors employed by listed companies, members of internal audit units affiliated with listed firms, or individuals working directly for the Exchange. Eligibility criteria should also be reported, including minimum professional experience, employment status, current internal-audit role, and whether participants were required to work in firms with formally established internal audit units.

The construct specification of internal audit effectiveness requires greater theoretical and statistical justification. The manuscript models “organizational effectiveness (OE), procedural effectiveness (PE), and relational effectiveness (CE)” as dimensions of internal audit effectiveness (IAE), with very high loadings of .977, .980, and .991. The authors should clearly state whether IAE is specified as a reflective-reflective higher-order construct or according to another hierarchical component model. The method used to estimate this higher-order construct in SmartPLS—such as repeated indicators, two-stage approach, or disjoint two-stage approach—should be reported because the interpretation of the model depends substantially on correct hierarchical construct specification.

The discriminant-validity results in Table 4 raise a major concern that must be resolved before the substantive findings can be accepted. The reported correlations among IAE, OE, PE, and CE are extremely high: IAE–OE = .990, IAE–PE = .981, IAE–CE = .975, and OE–PE = .969. These correlations are higher than several corresponding diagonal square-root-of-AVE values, which appears to violate the Fornell–Larcker criterion rather than support it. Therefore, the statement that “the model demonstrated discriminant validity at the construct level according to the Fornell–Larcker criterion” is not supported by the presented matrix. The authors should recheck Table 4, reassess the higher-order measurement model, report HTMT ratios, and determine whether these dimensions are empirically distinguishable.

Authors revised the manuscript and uploaded the new document.

1.2. Reviewer 2

Reviewer:

The sampling procedure requires considerably more methodological detail. The manuscript reports that “250 questionnaires were randomly distributed among eligible internal auditors” and that 195 usable questionnaires were returned. However, it is not explained how a random sampling procedure was possible when the authors simultaneously state that no comprehensive statistical frame of internal auditors was available. The authors should identify the sampling frame, describe how potential respondents were accessed, explain the randomization process, and report whether sampling occurred at the individual or company level. If probability sampling was not actually feasible, the method should be accurately described as convenience, purposive, network-based, or another appropriate non-probability sampling approach.

The sample-size calculation needs to be revised and justified more rigorously. The study reports the use of Cochran’s formula with “a standard deviation of 0.683 obtained from a preliminary study of 30 participants” and a margin of error of 0.10, yielding a minimum sample of 179. The authors should explain which variable was used to estimate the standard deviation of 0.683 and why a continuous-variable version of Cochran’s formula was preferred. More importantly, because the principal analysis is PLS-SEM, an SEM-specific power analysis would be more appropriate. The authors should report a statistical power calculation based on the maximum number of predictors of an endogenous construct, expected effect size, α level, and desired power, rather than relying exclusively on a generic survey sample-size formula.

The data-collection period should be reported unambiguously in Gregorian terms. The manuscript states that questionnaires were distributed “during the second half of the 2020–2021 period,” apparently based on the original expression referring to the

second half of 1399. This conversion is not sufficiently precise because the Persian calendar period spans parts of two Gregorian years. The authors should provide exact or approximate Gregorian months, for example “September 2020 to March 2021,” if that accurately reflects the original data collection period. Precise reporting is especially important because organizational conditions and audit practices may have been affected by contemporaneous economic or pandemic-related circumstances.

The measurement section should provide stronger evidence of instrument adaptation and validity. For example, the manuscript states that the Parker self-efficacy instrument, Snyder hope measure, Block and Kremen resilience scale, and Rahmani and Ghashghaei internal audit effectiveness questionnaire were used. However, the procedures used to adapt these instruments to the population of Iranian internal auditors are unclear. The authors should report whether validated Persian versions existed, whether translation and back-translation were undertaken, whether wording was modified to fit the audit context, how many experts evaluated content validity, and whether any quantitative content-validity indices such as CVR or CVI were calculated.

There is a potentially important inconsistency in the response formats across the instruments that should be clarified. The general methods section states that “all questionnaire items were rated using five-point Likert-type response formats,” yet the self-efficacy scale is described as ranging from “complete lack of confidence” to “complete confidence,” whereas the other instruments use agreement anchors ranging from strongly disagree to strongly agree. These are not identical response formats. The authors should describe the response scale for each instrument separately and explain whether scores were standardized or otherwise harmonized before inclusion in the structural model.

The authors should directly address the risk of common method bias. The methods section explicitly acknowledges that “all predictor and criterion measures were completed by each participant during the same data-collection period and were therefore obtained from a single source at a single point in time.” This design creates a substantial possibility of common method variance, particularly because both the independent and dependent variables are perceptual self-reports. The manuscript should report procedural remedies used to reduce this bias and should conduct an appropriate statistical assessment, such as full-collinearity VIFs, a marker-variable approach, or another method appropriate to PLS-SEM. Simply acknowledging the single-source design is insufficient.

Authors revised the manuscript and uploaded the new document.

2. Revised

Editor’s decision after revisions: Accepted.

Editor in Chief’s decision: Accepted.