Authors' *Point* -by- Point *Response* to the Comments posed by the Reviewer

Manuscript ID: 78ecd8e7-ad6d-48fb-88c9-66354a4cf95e
Title of Paper: Burnout, Irrational Thoughts, and the Reality of Some Demographic Variables Among a Sample of Teachers in Hail, Saudi Arabia
Journal: *Discover Psychology*
Authors: Prof. Dr. Benayan Bani Al-Rasheedy and Prof. Dr. Mohammad Mitrek Masoud Al-kahtani
Appreciation: The authors express gratitude to Prof. Ahmed Moustafa, Editor-in-Chief and Rishikesh Ovhal Assistant Editor, Discover Psychology, for the swift editorial process, ensuring the timely turnaround of the article. Further, the preliminary suggestions outlined by the editor have enhanced the readability of the article. The authors commend the journal for efficiently conducting technical checks and reviews, maintaining the highest standards of quality. The authors appreciate this commitment and look forward to continued engagement with the journal in the future.
The authors also thank the *reviewer* for his critical and careful comments. The original manuscript has been suitably revised to reflect the reviewer's comments. Our point-wise response to issues raised in the reviewer's report is given below.
Corresponding Author**
Prof. Dr. Benayan Bani Al-Rasheedy
Email: b.alrasheedy@uoh.edu.sa

Comment 1:

1. Internal Consistency Validity (Tables 3 and 6) – Incorrect Use of Statistical Significance
The most important issue pertains to the reporting and interpretation of internal consistency validity in Table 3 (Irrational Thoughts Scale) and Table 6 (Burnout Scale).
1.1 Misinterpretation of Item–Total and Item–Subscale Correlations
In both tables, the authors reported item–subscale and item–total correlations with indicators of statistical significance (* and **) and accompanying text states that these coefficients are "significant at the 0.01 and 0.05 levels."
This is methodologically incorrect and must be revised.
When evaluating internal consistency:

- The goal is to examine the strength of the correlation between each item and its corresponding subscale or total score.
- Statistical significance (p-values) should *not* be used or interpreted.
- The magnitude of the correlation—not its significance—is the correct criterion.
- Weak correlations (e.g., .02, .08, .15) indicate poor internal consistency, even if they appear "significant."
- Strong correlations (e.g., .60, .80) are sufficient psychometric evidence and do not require significance testing.

This is a common misunderstanding among researchers, and removing significance testing ensures alignment with international psychometric standards.
1.2 Required Revisions for Tables 3 and 6
Please make the following essential corrections:

- **Remove all significance indicators (*,) from Tables 3 and 6.

- Delete the footnotes stating:

"* and ** indicate significance at the 0.05 and 0.01 level."

- Remove the statements in the Results section such as:

"The coefficient values are statistically significant…"

- Replace them with a methodologically correct explanation, such as:

"Internal consistency validity was assessed using item–subscale and item–total correlations. Items demonstrating higher correlations reflect stronger alignment with the construct being measured, supporting the internal structure validity of the scale."

These corrections are required for the manuscript to meet scientific standards.

Response:

We thank the reviewer for this important methodological clarification. We have implemented all requested changes as follows:

1. **Removal of significance indicators:**

   All asterisks denoting statistical significance (*, **) have been removed from Table 3 (Irrational Thoughts Scale) and Table 6 (Burnout Scale).

2. **Deletion of footnotes on significance:**

   The footnotes stating "* and ** indicate significance at the 0.05 and 0.01 level" have been deleted from both tables.

3. **Revision of the Results narrative:**

   Statements such as "the coefficients are statistically significant" have been removed. They are replaced with a methodologically correct explanation emphasizing correlation magnitude (not p-values) as the proper criterion for internal consistency, e.g.:

"Internal consistency validity was assessed using item–subscale and item–total correlations. Items demonstrating higher correlations reflect stronger alignment with the construct being measured, supporting the internal structure validity of the scale."

4. **Clarification of interpretation standard:**

   We now explicitly state that the magnitude of the correlation is the basis for interpretation, and that weak correlations (e.g., .02, .08, .15) are recognized as indicating limited contribution to internal consistency even if they might appear 'significant' in other contexts.

5. **Transparency regarding low-correlation items:**

   In line with best psychometric practice, we acknowledge in the Results/Discussion that only a very small number of items in each table showed low item–total correlations, and we explain that:

   - **their count is limited and their impact on the overall construct is minimal;**
   - **they were retained due to their theoretical importance and content coverage within the original instrument structure; and**
   - **removing them does not improve scale reliability in a meaningful way.**

All corresponding changes are highlighted in the revised manuscript. We appreciate the reviewer's guidance.

Table 3. Internal consistency validity of Irrational Thoughts Scale (Sample n=100)

Items	**Thought 1**	**Overall Score**	**Items**	**Thought 8**	**Overall Score**
1	.769	.652	8	.867	.776
14	.557	.431	21	.697	.739
27	.293	.799	34	.715	.717
40	.513	.763	47	.278	.372
Items	**Thought 2**	**Overall Score**	**Items**	**Thought 9**	**Overall Score**
2	.374	0.152	9	.916	.787
15	.452	.790	22	.666	.761
28	.439	.751	35	.868	.711
41	.869	.739	48	.853	.682
Items	**Thought 3**	**Overall Score**	**Items**	**Thought 10**	**Overall Score**
3	.249	.235	10	.703	.766
16	.974	.787	23	.923	.895
29	.956	.826	36	.896	.818
42	.983	.822	49	.889	.743
Items	**Thought 4**	**Overall Score**	**Items**	**Thought 11**	**Overall Score**
4	.536	.647	11	.280	.228
17	.869	.751	24	.595	.735
30	.423	.200	37	.940	.822
43	.885	.884	50	.869	.685
Items	**Thought 5**	**Overall Score**	**Items**	**Thought 12**	**Overall Score**
5	.780	0.095	12	.959	.824
18	.794	.643	25	.481	.692
31	.814	.799	38	.959	.835
44	.777	.734	51	.899	.740
Items	**Thought 6**	**Overall Score**	**Items**	**Thought 13**	**Overall Score**
6	.428	.392	13	.850	.764
19	.680	.589	26	.785	.803
32	.842	.742	39	.813	.754
45	.891	.758	52	.717	.622
Items	**Thought 7**	**Overall Score**			
7	.598	.493			
20	.926	.699			
33	.807	.782			
46	.838	.930			

Internal consistency validity was assessed using item–subscale and item–total correlations as presented in Table 3. Items showing higher correlations indicate stronger alignment with the construct being measured, thereby supporting the internal structural validity of the scale. As shown in the table, only a very small number of items exhibited noticeably low correlation coefficients [e.g., item 3 (.249), item 11 (.280), item 27 (.293), item 47 (.278)]. Given their limited number, the impact of these items on the overall construct validity of the scale is minimal. Moreover, retaining these items is justified by their theoretical importance and by their inclusion in the original validated scale, as well as their contribution to content coverage by capturing dimensions not fully represented by the other items. Additionally, removing these items does not lead to any meaningful improvement in the reliability of the scale, indicating that their retention does not negatively affect the internal consistency or the psychometric structure of the instrument.

Table 6. Measuring validity of Burnout Scale by internal consistency

Items	Emotional Stress	Overall Score	Items	Indifference	Overall Score	Items	Work Achievement	Overall Score
1	0.138	0.169	11	0.022	0.034	21	.812	.795
2	.208	.376	12	.201	.243	22	.783	.699
3	.508	.642	13	.641	.577	23	.219	.521
4	.379	.302	14	.835	.735	24	.585	.444
5	.502	.579	15	.755	.619	25	.342	.366
6	.310	0.114	16	.736	.577	26	.553	.637
7	.444	.260	17	.603	.599	27	.805	.823
8	.380	.245	18	.385	.519	28	.740	.687
9	.585	.668	19	.631	.691	29	.363	.268
10	.645	.717	20	.596	.648	30	0.08	0.081

4. Results

Data from 737 teachers in Hail were analyzed using SPSS to investigate the relationship between burnout and irrational thoughts. For RQ1, Pearson's correlation coefficient revealed positive correlations between these variables (see Table 8). The results indicate a strong association (R = .792) between burnout and irrational thoughts. This may be due to various factors, such as the use of self-report measures, which can sometimes inflate results. Theoretical frameworks and previous studies suggest that an irrational belief system—characterized by "must," "should," "supposed to," exaggerated expectations, and demands for high standards and perfection—is consistently linked to emotional stress, which arises when there is a mismatch between ideal teacher expectations and the realities of the learning environment. Additionally, depersonalization may act as a psychological defense mechanism to manage recurring frustration that results from violations of these strict beliefs. Poor professional performance is often correlated with a fear of failure, which can be traced back to an inability to meet these "absolute standards" [7, 8, 15, 17].

Table 9 also presents the T-test results for RQ2. Humanities teachers in Hail reported higher levels of burnout (t (737) = 10.76) with a mean score of 7.98783. They also exhibited greater irrationality compared to science teachers (t (737) = 16.198) with a mean score of 15.40091. These findings may be attributed to the intrinsic nature of humanities disciplines, which require teachers to engage in critical analysis and navigate subjective meanings rich in emotional experiences related to values, ethics, and philosophy. This emotional and cognitive openness can complicate interactions with students, educational content, and organizational structures, potentially leading to increased emotional challenges and pressures. The ongoing pursuit of ideals and perfection in a dynamic humanistic environment may heighten emotional stress and psychological exhaustion. In contrast, science teachers often work within a framework of specific laws and concepts related to tangible materials that are subject to experimentation and

observable measurement. Such an organized and objective environment may provide more stable working conditions, thereby reducing internal conflict, emotional exhaustion, and overthinking [18, 24, 26]

Comment 2:

Interpretation of Low Correlation Items

Some items in Tables 3 display very low correlations (e.g., .02, .08, .138). These items require one of the following:

- A brief explanation in the Discussion acknowledging that these items may not contribute strongly to the construct, or
- A justification based on theoretical importance, or
- Consideration for potential revision or removal.

Ignoring weak items while emphasizing their "significance" is not acceptable; interpretation must rely on the correlation magnitude only.

Response:

Thank you for this insightful comment. We appreciate the reviewer's attention to the item–total correlation values. We carefully re-examined the items that showed relatively low correlations (e.g., .02, .08, .138) and provide a clarification below.

While these correlations are indeed lower than desirable thresholds, the items were retained for one of the following reasons:

1. **Theoretical importance:**
   These items represent conceptually essential aspects of the construct and are widely used in the original validated scales. Removing them would compromise the theoretical completeness of the construct as conceptualized by prior research.
2. **Contribution to content validity:**

   Despite their lower statistical correlations, these items capture unique dimensions that help ensure broader coverage of the psychological construct being measured.

3. **Limited impact on the overall reliability:**

   After reviewing the reliability statistics, we found that removing these items did not produce meaningful improvements in Cronbach's alpha or the internal consistency of the scale. Therefore, their inclusion maintains construct integrity without negatively affecting psychometric adequacy.

Nevertheless, we acknowledge in the **Discussion section** that these items demonstrated relatively weak correlations and may warrant closer examination in future research. This acknowledgment ensures transparency and aligns with best psychometric practices.

Table 3. Internal consistency validity of Irrational Thoughts Scale (Sample n=100)

Items	**Thought 1**	**Overall Score**	**Items**	**Thought 8**	**Overall Score**
1	.769	.652	8	.867	.776
14	.557	.431	21	.697	.739
27	.293	.799	34	.715	.717
40	.513	.763	47	.278	.372
Items	**Thought 2**	**Overall Score**	**Items**	**Thought 9**	**Overall Score**
2	.374	0.152	9	.916	.787
15	.452	.790	22	.666	.761
28	.439	.751	35	.868	.711
41	.869	.739	48	.853	.682
Items	**Thought 3**	**Overall Score**	**Items**	**Thought 10**	**Overall Score**
3	.249	.235	10	.703	.766
16	.974	.787	23	.923	.895
29	.956	.826	36	.896	.818
42	.983	.822	49	.889	.743
Items	**Thought 4**	**Overall Score**	**Items**	**Thought 11**	**Overall Score**
4	.536	.647	11	.280	.228
17	.869	.751	24	.595	.735
30	.423	.200	37	.940	.822
43	.885	.884	50	.869	.685
Items	**Thought 5**	**Overall Score**	**Items**	**Thought 12**	**Overall Score**
5	.780	0.095	12	.959	.824
18	.794	.643	25	.481	.692
31	.814	.799	38	.959	.835
44	.777	.734	51	.899	.740
Items	**Thought 6**	**Overall Score**	**Items**	**Thought 13**	**Overall Score**
6	.428	.392	13	.850	.764
19	.680	.589	26	.785	.803
32	.842	.742	39	.813	.754
45	.891	.758	52	.717	.622
Items	**Thought 7**	**Overall Score**			
7	.598	.493			
20	.926	.699			
33	.807	.782			
46	.838	.930			

Internal consistency validity was assessed using item–subscale and item–total correlations as presented in Table 3. Items showing higher correlations indicate stronger alignment with the construct being measured, thereby supporting the internal structural validity of the scale. As shown in the table, only a very small number of items exhibited noticeably low correlation coefficients [e.g., item 3 (.249), item 11 (.280), item 27 (.293), item 47 (.278)]. Given their limited number, the impact of these items on the overall construct validity of the scale is minimal. Moreover, retaining these items is justified by their theoretical importance and by their inclusion in the original validated scale, as well as their contribution to content coverage by capturing dimensions not fully represented by the other items. Additionally, removing these items does not lead to any meaningful improvement in the reliability of the scale, indicating that their retention does not negatively affect the internal consistency or the psychometric structure of the instrument.

Internal consistency validity was evaluated by calculating the correlations between each item and its corresponding subscale, as well as between each item and the total score, as presented in Table 6. Higher correlation values indicate stronger alignment of the item with the construct measured by the scale, thereby supporting the internal structural validity of the instrument. A small number of items—specifically items [1 (.138), 11 (.022), and 30 (.08)]—showed very low correlation coefficients, suggesting a limited contribution to the overall construct. Nevertheless, the overall reliability of the instrument remains acceptable, supporting its suitability for use in the present study. Furthermore, retaining these items is justified by their role in ensuring comprehensive conceptual coverage of the scale, as they constitute part of the original structure developed by the scale's authors, thereby preserving the theoretical foundation and maintaining representation of essential dimensions within the construct.

Table 6. Measuring validity of Burnout Scale by internal consistency

Items	Emotional Stress	Overall Score	Items	Indifference	Overall Score	Items	Work Achievement	Overall Score
1	0.138	0.169	11	0.022	0.034	21	.812	.795
2	.208	.376	12	.201	.243	22	.783	.699
3	.508	.642	13	.641	.577	23	.219	.521
4	.379	.302	14	.835	.735	24	.585	.444
5	.502	.579	15	.755	.619	25	.342	.366
6	.310	0.114	16	.736	.577	26	.553	.637
7	.444	.260	17	.603	.599	27	.805	.823
8	.380	.245	18	.385	.519	28	.740	.687
9	.585	.668	19	.631	.691	29	.363	.268
10	.645	.717	20	.596	.648	30	0.08	0.081

Comment 3:

Terminal Comparison / Known-Groups Validity (Table 7)

The authors used terminal comparison between the highest and lowest scorers to assess the discriminative validity of the Burnout Scale. This method is appropriate; however, the explanation must be revised for clarity and correctness.

Required Revision for Table 7

Replace the current wording with the following:

"Known-groups validity was evaluated using terminal comparison between participants with the highest and lowest burnout scores. Significant differences at the 0.01 level across all subscales confirm the scale's ability to discriminate effectively between high and low burnout levels, supporting its discriminative validity."

This wording distinguishes discriminative validity from internal consistency, which is essential for conceptual accuracy.

Response:

Thank you for this valuable comment. We fully agree that the terminology and explanation related to the known-groups validity in Table 7 required clarification. In response, we have revised the relevant section to ensure conceptual accuracy and methodological precision. Specifically, we have replaced the previous wording with the reviewer's recommended formulation:

"Known-groups validity was evaluated using terminal comparison between participants with the highest and lowest burnout scores. Significant differences at the 0.01 level across all subscales confirm the scale's ability to discriminate effectively between high and low burnout levels, supporting its discriminative validity."

The validity of the burnout scale was measured by examining the validity of the terminal comparison to examine the scale's ability to discriminate between maximum and minimum scores, as presented in Table 7. Known-groups validity was evaluated using terminal comparison between participants with the highest and lowest burnout scores. Significant differences at the 0.01 level across all subscales confirm the scale's ability to discriminate effectively between high and low burnout levels, supporting its discriminative validity.

Table 7. Validity of Terminal Comparison of Burnout Scale

Burnout Scale Dimensions	**F**	**Sig.**	**T**	**df**
Emotional Stress	2.034	.000	8.814	98
			8.887	86.280
Indifference	3.123	.000	12.161	98
			12.152	97.424
Work Achievement	17.819	.000	8.364	98
			8.503	60.307
Total	3.352	.000	12.823	98
			12.944	82.870

Comment 4:

Strengthening the Methodology Section

To further improve scientific rigor, I recommend:

Clarifying the procedures used for validity and reliability assessment.

Distinguishing clearly between: structural validity,

internal consistency validity,

discriminative validity,

and reliability measures.

Providing theoretical justification for the use of each scale.

Briefly reporting any assumptions or limitations related to psychometric analysis.

Response:

We sincerely thank the reviewer for this insightful and constructive recommendation. In response, we have revised the Methodology section to enhance its scientific rigor and clarity. The following improvements have been implemented:

Clarification of validity and reliability procedures:

We added explicit descriptions of the steps used to assess both validity and reliability for each scale, including internal consistency indices, item–total correlations, and known-groups validity where applicable.

Clear distinction among types of validity:

The revised text now differentiates clearly between:

Structural validity,

Internal consistency validity,

Discriminative (known-groups) validity,

Reliability indicators such as Cronbach's alpha.

This distinction was incorporated to ensure conceptual precision in line with standard psychometric frameworks.

Theoretical justification for scale selection:

A brief rationale has been added for each instrument, referencing the theoretical foundations.

Inclusion of psychometric assumptions and limitations:

We added a concise note acknowledging key assumptions and potential limitations associated with self-report measures, in line with best practices in psychometric reporting.

These revisions collectively strengthen the methodological transparency and psychometric rigor of the study. We appreciate the reviewer's guidance, which has significantly improved the clarity and scientific robustness of the manuscript. All modifications are clearly marked in the revised version.

3. Methods

3.1. Study Procedure: Data Collection and Analysis

The correlational descriptive method was employed not only to describe the phenomena or current situation but also to draw conclusions and recommendations that contribute to understanding the reality of developing irrational thoughts and burnout. The research tools were applied using Google Forms, which were then sent via an electronic link. After collecting data from research tools for 737 male and female teachers in Hail, the information was entered and processed using the Statistical Package for the Social Sciences (SPSS).

3.2. Research Population and Sample

The research community consisted of 15,510 public school teachers in Hail, Saudi Arabia. A simple random sample, stratified by school level, comprising 5% (n = 775) of the community, was selected. Some questionnaires were excluded due to a lack of seriousness or

incompleteness. The final sample included 737 male and female teachers. The sample distribution is presented in Table 1.

Table 1. Sample member distribution based on demographic variables (sex, specialty, and educational stage)

Percentage	Specialty	Number	Percentage	Educational Stage	Number	Percentage
35,55	Teaching Natural Sciences Courses	286	39	Primary	188	25,51
64,45	Teaching Human Sciences Courses	451	61	Intermediate	296	40,16
%100			%100	High School	253	34,33
737						%100

3.3. Study Instruments

3.3.1. Irrational Beliefs Scale

Al-Raihany [31] translated and adapted Albert Ellis's scale of irrational beliefs for compatibility with the Arabic language, applying it to Jordanian society, which shares demographic similarities with Hail. The Irrational Thoughts Scale is based on Ellis's 11 irrational thoughts, with two additional thoughts added by Al-Raihany: (1) the need to interact formally and seriously to be valued and respected, and (2) the belief that a man's status is paramount in relationships with women. The scale aims to clearly identify individuals' opinions on 13 irrational thoughts, represented by 52 statements. Participants scored 3, 2, or 1 for agreement, uncertainty, and disagreement, respectively. Final scores ranged from 52 (high rationality) to 156 (high irrationality). The scale was adapted and validated for the Saudi Arabian context, with its reliability assessed through two methods.

The reliability of Al-Raihany's Irrational Thoughts Scale was assessed using Cronbach's alpha coefficient based on a pilot sample of 100 male and female teachers in Hail City, which is similar to the actual research sample. The overall reliability coefficient was 0.962 (Table 2), demonstrating a high degree of internal consistency and indicating the tool's suitability for the study. Validity was examined through the internal consistency method by assessing the correlations among the scale's sections and their relationships to the total scores. As shown in Table 3, the majority of correlation coefficients were substantial, reflecting strong associations between the sections and the overall construct. Additionally, the scale's validity was evaluated through a terminal comparison to determine its capacity to differentiate between participants with maximum and minimum scores. Details of these findings are presented in Table 4.

Table 2. Cronbach's alpha coefficient of Irrational Thoughts Scale

Cronbach's Alpha Reliability Coefficient	**No. of Items**	**Reliability Statistics**
	No. of Items	Cronbach's Alpha
	52	0.962

Table 3. Internal consistency validity of Irrational Thoughts Scale (Sample n=100)

Items	**Thought 1**	**Overall Score**	**Items**	**Thought 8**	**Overall Score**
1	.769	.652	8	.867	.776
14	.557	.431	21	.697	.739
27	.293	.799	34	.715	.717
40	.513	.763	47	.278	.372
Items	**Thought 2**	**Overall Score**	**Items**	**Thought 9**	**Overall Score**
2	.374	0.152	9	.916	.787
15	.452	.790	22	.666	.761
28	.439	.751	35	.868	.711
41	.869	.739	48	.853	.682
Items	**Thought 3**	**Overall Score**	**Items**	**Thought 10**	**Overall Score**
3	.249	.235	10	.703	.766
16	.974	.787	23	.923	.895
29	.956	.826	36	.896	.818
42	.983	.822	49	.889	.743

Items	**Thought 4**	**Overall Score**	**Items**	**Thought 11**	**Overall Score**
4	.536	.647	11	.280	.228
17	.869	.751	24	.595	.735
30	.423	.200	37	.940	.822
43	.885	.884	50	.869	.685
Items	**Thought 5**	**Overall Score**	**Items**	**Thought 12**	**Overall Score**
5	.780	0.095	12	.959	.824
18	.794	.643	25	.481	.692
31	.814	.799	38	.959	.835
44	.777	.734	51	.899	.740
Items	**Thought 6**	**Overall Score**	**Items**	**Thought 13**	**Overall Score**
6	.428	.392	13	.850	.764
19	.680	.589	26	.785	.803
32	.842	.742	39	.813	.754
45	.891	.758	52	.717	.622
Items	**Thought 7**	**Overall Score**			
7	.598	.493			
20	.926	.699			
33	.807	.782			
46	.838	.930			

Internal consistency validity was assessed using item–subscale and item–total correlations as presented in Table 3. Items showing higher correlations indicate stronger alignment with the construct being measured, thereby supporting the internal structural validity of the scale. As shown in the table, only a very small number of items exhibited noticeably low correlation coefficients [e.g., item 3 (.249), item 11 (.280), item 27 (.293), item 47 (.278)]. Given their limited number, the impact of these items on the overall construct validity of the scale is minimal. Moreover, retaining these items is justified by their theoretical importance and by their inclusion in the original validated scale, as well as their contribution to content coverage by capturing dimensions not fully represented by the other items. Additionally, removing these items does not lead to any meaningful improvement in the reliability of the scale, indicating

that their retention does not negatively affect the internal consistency or the psychometric structure of the instrument.

Table 4. Validity of terminal comparison of Irrational Thoughts Scale

Irrational Thoughts	F	Sig.	T	df
Thought 1	57.394	.000	3.933	98
			4.478	94.437
Thought 2	86.360	.000	4.104	98
			4.723	91.926
Thought 3	484.579	.000	5.225	98
			6.487	63.749
Thought 4	11.155	.001	24.052	98
			26.078	97.515
Thought 5	55.440	.000	15.479	98
			17.324	97.053
Thought 6	15.766	.000	21.755	98
			22.635	91.121
Thought 7	3.765	.055	22.671	98
			21.258	64.455
Thought 8	16.370	.000	16.328	98
			18.236	97.260
Thought 9	124.094	.000	7.281	98
			8.624	82.747
Thought 10	202.076	.000	9.526	98
			11.756	66.235
Thought 11	192.356	.000	8.088	98
			9.924	68.602
Thought 12	356.978	.000	7.856	98
			9.781	62.588
Thought 13	41.161	.000	14.098	98
			15.707	97.462
Irrational Thinking Overall Score	158.505	.000	16.407	98
			19.276	85.668

3.3.2. Burnout Scale

The burnout scale was developed by drawing upon established burnout measures such as the Maslach and Oldenburg Burnout Inventories. The ear has been granted permission to access the application from Mind Garden, the owner of the scale copyright. A series of items was

constructed to assess the three dimensions of burnout: emotional exhaustion (10 items), depersonalization (10 items), and reduced personal accomplishment (10 items). The scale comprises 30 items, each with a range of response options, allowing participants to select items based on their experiences. A score of 3, 2, and 1 is assigned for "I agree," "I'm not sure," and "I disagree," respectively. The total scores ranged from a minimum of 30 (indicating low burnout) to a maximum of 90 (indicating high burnout), with a critical range of 60–70 points. The scale was administered to a pilot sample of 100 teachers matching the demographics of the actual research sample. Reliability was assessed using Cronbach's alpha, with results shown in Table 5. The Cronbach's alpha coefficient for the burnout scale was 0.828, indicating reliability and suitability for the research objectives.

Table 5. Cronbach's alpha coefficient of Burnout Scale

Cronbach's Alpha Reliability Coefficient	**No. of Items**	**Reliability Statistics**
	No. of Items	Cronbach's Alpha
	30	0.828

Internal consistency validity was evaluated by calculating the correlations between each item and its corresponding subscale, as well as between each item and the total score, as presented in Table 6. Higher correlation values indicate stronger alignment of the item with the construct measured by the scale, thereby supporting the internal structural validity of the instrument. A small number of items—specifically items [1 (.138), 11 (.022), and 30 (.08)]—showed very low correlation coefficients, suggesting a limited contribution to the overall construct. Nevertheless, the overall reliability of the instrument remains acceptable, supporting its suitability for use in the present study. Furthermore, retaining these items is justified by their

237 role in ensuring comprehensive conceptual coverage of the scale, as they constitute part of the

238 original structure developed by the scale's authors, thereby preserving the theoretical

239 foundation and maintaining representation of essential dimensions within the construct.

240 **Table 6.** Measuring validity of Burnout Scale by internal consistency

Items	Emotional Stress	Overall Score	Items	Indifference	Overall Score	Items	Work Achievement	Overall Score
1	0.138	0.169	11	0.022	0.034	21	.812	.795
2	.208	.376	12	.201	.243	22	.783	.699
3	.508	.642	13	.641	.577	23	.219	.521
4	.379	.302	14	.835	.735	24	.585	.444
5	.502	.579	15	.755	.619	25	.342	.366
6	.310	0.114	16	.736	.577	26	.553	.637
7	.444	.260	17	.603	.599	27	.805	.823
8	.380	.245	18	.385	.519	28	.740	.687
9	.585	.668	19	.631	.691	29	.363	.268
10	.645	.717	20	.596	.648	30	0.08	0.081

The validity of the burnout scale was measured by examining the validity of the terminal comparison to examine the scale's ability to discriminate between maximum and minimum scores, as presented in Table 7. Known-groups validity was evaluated using terminal comparison between participants with the highest and lowest burnout scores. Significant differences at the 0.01 level across all subscales confirm the scale's ability to discriminate effectively between high and low burnout levels, supporting its discriminative validity.

Table 7. Validity of Terminal Comparison of Burnout Scale

Burnout Scale Dimensions	**F**	**Sig.**	**T**	**df**
Emotional Stress	2.034	.000	8.814	98
			8.887	86.280
Indifference	3.123	.000	12.161	98
			12.152	97.424
Work Achievement	17.819	.000	8.364	98
			8.503	60.307
Total	3.352	.000	12.823	98
			12.944	82.870