Available online at website: https://jurnal.iaii.or.id/index.php/RESTI

JURNAL RESTI
(Rekayasa Sistem dan Teknologi Informasi)
Vol. 10 No. 3 (2026) 712 - 722 e-ISSN: 2580-0760

Fuzzy-IOWA-Based Group Decision Support System (GDSS) for Interpreting DASS Scores with Dynamic Expert Weighting

Wiharto[1], Eka P. Meravigliosi[2], Umi Salamah[3], Esti Suryani[4], Vihi Atina[5], Pradityo Utomo[6]
[1,2,3]Department of Informatics, Universitas Sebelas Maret, Surakarta, Indonesia
[4]Department of Data Science, Universitas Sebelas Maret, Surakarta, Indonesia
[5]Department of Software Engineering Technology, Universitas Duta Bangsa, Surakarta, Indonesia
[6]Department of Informatic Management, Universitas Merdeka Madiun, Madiun, Indonesia

[*]*wiharto@staff.uns.ac.id*

Abstract

This research develops a Group Decision Support System (GDSS) to address subjectivity and ambiguity in the interpretation of Depression Anxiety Stress Scales (DASS-21) scores, particularly in cases of symptom overlap across dimensions. The system introduces a structured weighting mechanism in which the influence of each expert is determined based on objective criteria, including clinical experience, education level, and academic contributions. The methodology applies Simple Additive Weighting (SAW) to quantify the relative importance of five clinical experts, resulting in Expert 2 (27.66%) and Expert 1 (27.36%) having the highest influence within the group. These weights are then incorporated into a Fuzzy-Induced Ordered Weighted Averaging (Fuzzy-IOWA) framework to aggregate expert judgments into a unified consensus model. The results indicate that the proposed approach is able to produce a consistent interpretation structure, with a tendency toward the Anxiety dimension in cases of overlapping symptoms. By integrating expert consensus with patient self-report scores, the system generates a structured interpretation profile of DASS-21 responses. The proposed GDSS provides a systematic and transparent aggregation framework that reflects expert reasoning. However, it is intended as a methodological support tool rather than a substitute for clinical diagnosis. The novelty of this work lies in the integration of SAW-based objective expert capability quantification with Fuzzy-IOWA aggregation and QGDD-based consensus characterization, forming a transparent and mathematically accountable GDSS pipeline that has not been previously applied in the context of DASS-21 interpretation.

Keywords: DASS; expert weighting; Fuzzy-IOWA; GDSS; mental health

How to Cite: Wiharto, E. P. Meravigliosi, U. Salamah, E. Suryani, V. Atina, and P. utomo, "Fuzzy-IOWA-Based Group Decision Support System (GDSS) for Interpreting DASS Scores with Dynamic Expert Weighting", *J. RESTI (Rekayasa Sist. Teknol. Inf.)* , vol. 10, no. 3, pp. 712 - 722, Jun. 2026.
*Permalink/*DOI: https://doi.org/10.29207/resti.v10i3.7340

Received: November 16, 2025
Accepted: June 1, 2026
Available Online: June 18, 2026

This is an open-access article under the CC BY 4.0 License
Published by Ikatan Ahli Informatika Indonesia

1. Introduction

The global burden of mental health disorders, manifested primarily through depression, anxiety, and stress, has become one of the most significant and pressing public health challenges in the contemporary era. The scale of this problem is comprehensively documented in global reports from the World Health Organization (WHO), which highlight its significant impact on individual quality of life and healthcare systems as a whole. In line with this alarming global trend, the situation in Indonesia also shows a similar urgency, with studies highlighting a significant need to develop assessment methods that are accurate, efficient, and widely accessible.

In psychological assessment practice, the Depression Anxiety Stress Scales (DASS) framework plays a vital role. The DASS instrument itself was initially developed to provide a valid psychometric measurement of three distinguishable negative emotional states, depression, anxiety, and stress. Which

often overlap in clinical practice [1]. The DASS instrument itself has demonstrated robust psychometric properties across various populations. For instance, studies using confirmatory factor analysis have repeatedly confirmed the validity of its three-factor structure, reinforcing its utility in distinguishing these emotional states, particularly in clinical patient samples. The fundamental strength of the DASS framework lies in its solid theoretical foundation, namely the Tripartite Model, which characterizes negative moods related to depression and anxiety through three main components: symptoms of increased negative affect or general distress, low levels of positive affect, and physiological hyperarousal [2]. Further support for the DASS-21 in mapping the Tripartite Model was also shown in research on clinical patients, which confirmed DASS's ability to distinguish the constructs of 'anhedonia' (as a unique component of depression) and 'physiological hyperarousal' (as a unique component of anxiety) [3].

Nevertheless, interpreting complex DASS score profiles remains challenging, especially in comorbid cases where diagnostic decisions depend on clinicians' expertise. In Group Decision-Making (GDM), this is viewed as a multi-scale decision problem, in which experts use different preference scales or information granularities [4].

To bridge the gap between clinical ambiguity and systematic assessment, the initial wave of research utilized Machine Learning (ML) to classify mental conditions based on DASS score patterns [5] [6]. For instance, supervised models such as Random Forest have been applied to predict the severity of depression, anxiety, and stress from DASS-21 data [7]. While these approaches offer high predictive accuracy, their 'black box' nature often limits clinical trust, driving a shift toward Explainable AI (XAI) [8] and knowledge-based Clinical Decision Support Systems (CDSS) to enhance transparency and accountability [9]. Within this context, Fuzzy Logic has emerged as a well-suited methodology for managing the qualitative uncertainty and linguistic ambiguity inherent in clinical assessments [10]. However, to further improve decision validity and address differing opinions among practitioners, the focus has evolved toward more sophisticated Group Decision Support Systems (GDSS) that can aggregate multiple expert perspectives.

The implementation of GDSS, however, introduces the challenge of determining the relative importance of each expert's opinion within a group. Standard aggregation methods, such as Ordered Weighted Averaging (OWA), often ignore the source of the opinion by focusing only on the rank of the values [11]. This is suboptimal in clinical settings where a senior expert's judgment should carry more weight than that of a junior practitioner. To resolve this, this study specifically adopts the Induced Ordered Weighted Averaging (IOWA) operator [12]. By using each expert's objective capability score as an external inducing value, this research ensures that professional credibility quantitatively determines influence in the final decision [13]. This approach aligns with the need for an effective GDSS model that can combine vague linguistic judgments fairly according to respective levels of expertise [14].

The Fuzzy-IOWA approach allows expert judgments to be systematically integrated into a computational decision model. Basic research in this field has already demonstrated that the GDSS approach is indeed effective, specifically noting that the differences in priorities between GDSS and standard DASS models highlight a system's inherent ability to adjust its focus based on the individual interpretations and nuanced preferences of each psychologist [15]. This finding underscores a secondary and more complex challenge, which is the fact that if valid expert preferences can differ and their capability levels are not uniform, then a more sophisticated mechanism is required to combine these views optimally. Consequently, reliance on simple consensus or unweighted averaging is insufficient when the credibility and professional background of the information sources vary significantly within a professional panel.

Therefore, this research proposes a significant innovation in the form of a GDSS model based on Fuzzy-IOWA that incorporates a measurable and objective expert weighting mechanism. This model systematically calculates the influence of each professional by utilizing specific capability criteria, such as clinical experience, patient volume, educational qualifications, and scientific publications. Unlike previous models that might overlook the seniority of the source, these resulting weights are used directly in the Fuzzy-IOWA aggregation process to produce a consensus result that is both transparent and academically accountable. This structured approach ensures that an expert's influence on the final diagnostic interpretation is proportional to their clearly quantified level of expertise, thereby attempting to bridge the gap between subjective clinical intuition and computational assessment. Specifically, the novelty of this work is the integration of three components not previously combined in DASS-21 interpretation frameworks: (1) SAW-based objective expert capability quantification as inducing variables, (2) Fuzzy-IOWA aggregation that prioritizes inputs proportionally to quantified expertise, and (3) QGDD-based characterization of consensus tendency, forming an end-to-end pipeline that is both transparent and clinically grounded.

2. Methods

The research methodology was systematically designed to achieve the research objectives. This section describes the research framework that guided the workflow, the data collection methods, and the data analysis stages conducted chronologically and in detail, presenting the mathematical formulas used.

2.1 Research Framework

As a guiding blueprint for the investigation, the research framework is executed through a series of integrated and highly structured phases, as visually synthesized in Figure 1.

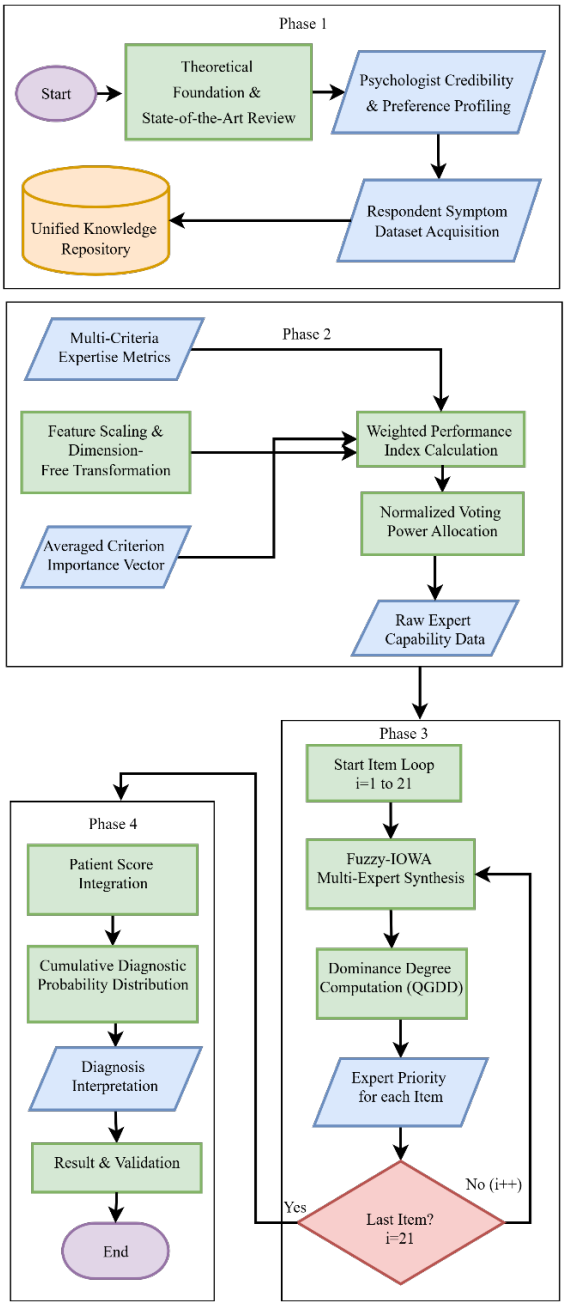


Figure 1. Research Framework Flowchart

This systematic progression is designed to maintain logical consistency throughout the study. The workflow is initiated with a critical problem identification stage, which serves to define the boundaries of the clinical ambiguity being addressed. This is followed by a robust data collection phase to gather both expert credentials and patient metrics. Subsequently, the process moves into the implementation of the computational model, where the SAW and Fuzzy-IOWA algorithms are synthesized to process the acquired data. Finally, the framework culminates in a detailed results analysis and the derivation of meaningful conclusions that bridge the gap between theoretical modeling and practical diagnostic application.

2.2 Data Collection Methods

This research utilizes secondary data collected from various sources. First, a Literature Study was conducted by gathering theories and methods from scientific journals, reports, and textbooks related to DASS, GDSS, Fuzzy Logic, and IOWA. Second, Expert Capability and Preference Data were collected as primary data through a field study conducted at the Psychology Services Unit of Universitas Sebelas Maret.

This process involved interviews and questionnaires with five psychology experts to obtain three types of data: (a) capability data including experience, education level, and other track records, (b) their ordinal preference data for each DASS item as a reference for determining disorder priorities, and (c) percentage data for each expert capability parameter to find the average weight from the five experts. Third, Patient Sample Data, a set of DASS-21 questionnaire answers from a sample patient, was defined to test the system's workflow and functionality.

2.3 Model Implementation Stages

This stage is the core of the research, where the model is computationally designed and implemented. Each step is explained with the mathematical formulas that form its basis.

Within the Group Decision Support System (GDSS) framework, not all experts possess an equal level of expertise or credibility. An expert's credibility cannot be represented by a single factor, but rather by multiple aspects such as qualifications and experience [16]. Therefore, a mechanism is required to objectively quantify each expert's capability, the results of which are then used to determine their importance weight in the decision aggregation process. This process is conducted through two main stages: (1) identification and determination of relevant capability criteria based on theoretical foundations and empirical studies, and (2) determination of initial weights for each criterion reflecting its relative importance in the context of clinical diagnostic assessment. This approach aligns with the 'objective weight' determination method for experts in Group Decision-Making (GDM), which determines weights based on actual data (like opinions or credibility) rather than just 'subjective weights' like reputation or status [17]. The expert capability data for each criterion is presented in Table 1.

The selection of criteria is based on parameters widely recognized in psychological and cognitive literature as indicators of expertise, as follows. Clinical Experience (Years of Experience and Number of Patients): Expert skills, especially in clinical fields, develop through stages heavily dependent on experience. The Dreyfus model explains that a practitioner transitions from a

novice level, rigidly following rules, to an expert level, capable of making holistic and intuitive decisions based on extensive experience [18]. To operationalize this concept, this study uses two parameters: Years of Experience, to measure the duration or length of experience, and Number of Patients to measure the volume or density of that experience. The use of "years of experience" is a common approach in psychology studies to represent an expert's quality [19]. Although 'clinical experience' is often used as a proxy for 'expertise', research also shows that experience alone may not be sufficient to produce expertise in psychological assessment, reinforcing the need to measure it alongside other criteria [20] such as educational qualifications and publication count. Formal Qualifications: Academic qualification is the foundation of an expert's credibility. This parameter represents the mastery of theoretical frameworks and evidence-based knowledge that underpins clinical practice. Related studies mentioned earlier [19] also show that an expert's credentials (like professional background and education) are significant factors influencing the audience's perception of their expertise. Other research also confirms that the level of academic training significantly affects the types and methods of assessment used by psychologists, indicating that formal training is a distinct component of expertise from mere clinical experience [20]. Academic Contributions (Number of Publications): Involvement in research and scientific publication serves as a differentiator for experts who not only apply knowledge but also contribute to its development. This signifies a deep understanding of research methodology and a commitment to staying relevant with the latest scientific advancements, which adds "prestige" to their credentials [19]. The determination of initial weights for each criterion parameter (like 'Years of Experience' or 'Patients') is based on aggregating the collective views of the experts. This model collects the percentage of importance inputted by *each* expert for *each* criterion. These percentage values are then averaged to produce one raw weight score per criterion.

Table 1. Expert Capability Table

Psychologist	Years of Experience	Patient	Education	Publication
P1	8	68	S2	32
P2	6	650	S2	2
P3	2	200	S2	1
P4	3	350	S2	4
P5	4	300	S2	2

After the criteria and weights are determined, the next step is to normalize the raw data in Table 1. This process is a fundamental step in the Simple Additive Weighting (SAW) method, which requires the normalization of the decision matrix so that each attribute's rating becomes dimension-free and comparable [21]. Normalization aims to eliminate unit bias and transform all criteria values into a uniform range [0, 1]. This process is performed using Equation 1.

$$r_{ij} = \frac{C_{ij} - min(C_j)}{max(C_j) - min(C_j)} \quad (1)$$

r_{ij} is the normalized value for expert i on criterion j, C_{ij} is the actual value of expert i on criterion j, $min(C_j)$ is the lowest value from all experts for criterion j, and $max(C_j)$ is the highest value from all experts for criterion j.

The final capability score (V_i) for each expert is calculated by summing the multiplication results of each normalized value with its respective criterion weight. This weighted summation method (Equation 2) is known as the Simple Additive Weighting (SAW) method, the basic concept of which is to find the weighted sum of each alternative's performance rating on all attributes to obtain a total score [21]:

$$V_i = \sum_{j=1}^{n} (w_j \times r_{ij}) \quad (2)$$

V_i is the final capability score for expert i, w_j is the weight for criterion j (taken from the expert weight averaging process), r_{ij} is the normalized value of expert i for criterion j, and n is the total number of criteria (in this case, 4).

This stage is a crucial step to convert raw capability scores into proportional weights ready for use. The goal is to ensure that the total influence of all experts is 1 (or 100%), so each expert's weight represents their "portion" of influence in the group decision.

$$w_i = \frac{V_i}{\sum_{i=1}^{N} V_i} \quad (3)$$

As written in Equation 3, to determine the final weight or "voting power" (w_i) of each expert, the raw capability score (v_i) calculated in the previous stage is divided by the total sum of all experts' raw capability scores (ΣV_i) . This division process effectively converts the raw score into a proportional value representing the portion or percentage of contribution from each expert towards the overall group decision.

To ensure the reproducibility of the model, it is essential to clarify the data transformation process. Expert preferences are encoded using a nine-point linguistic scale, which is then transformed into Triangular Fuzzy Numbers (TFN) defined as (l,m,u), where l, m, and u denote the lower, middle, and upper bounds, respectively. The linguistic scale is defined as follows: “Very Low” (0.00, 0.00, 0.25), “Low” (0.00, 0.25, 0.50), “High” (0.25, 0.50, 0.75), and “Very High” (0.50, 0.75, 1.00). This mapping allows qualitative expert judgments to be represented numerically while preserving uncertainty.

After constructing the fuzzy preference matrix, defuzzification is performed using the Center of Area (CoA) method (Equation 4), defined as:

$$CoA = \frac{l+m+u}{3} \quad (4)$$

This process differs from ordinary weighted averaging because it involves two main components: the expert's capability score (V_i) as the order-inducing value and the OWA weight (W) as the determinant of the importance level. The IOWA aggregation process for each preference matrix element p_{ij} is as follows:

For each pair of alternatives (i, j), the preference values from all experts $(p_{ij1}, p_{ij2}, \dots, p_{ijn})$ are sorted based on their expert capability scores $(V_1, V_2, \dots, V_n)$ from highest to lowest. This process yields a new sorted preference vector, called $B_{ij} = (b_{ij1}, b_{ij2}, \dots b_{ijn})$, where b_{ij1} is the preference from the expert with the highest capability, which is subsequently aggregated during the weighting phase by multiplying the sorted preference vector b_{ij} with the expert capability weight vector that has been calculated in Equation 2.

Mathematically, this Fuzzy-IOWA aggregation process is formulated in Equation 5.

$$p_{ij}^c = F_W(B_{ij}) = \sum_{k=1}^{N} w_k \cdot b_{ijk} \quad (5)$$

p_{ij}^c is the collective preference value for alternative i against j , N is the total number of experts , w_k is the importance weight (IOWA Weight) of the k-th expert, and b_{ijk} is the preference value from the expert with the k-th rank for alternative i against j. With this approach, the preference matrix, which was initially qualitative, is transformed into a crisp value through defuzzification at the beginning, and then aggregated based on expert credibility, resulting in a single crisp matrix that represents the group consensus.

From the aggregated group preference matrix (P_k), the final score for each condition is calculated using the Quantified Guided Dominance Degree (QGDD) method. This method is used to measure the quantitative dominance level of one alternative over all other alternatives [22]. The QGDD score for each alternative i is calculated using Equation 6.

$$QGDD_i = F_Q(p_{ij}^c, j = 1, \dots, n, i \neq j) \quad (6)$$

The F_Q function in Equation (6) is an OWA operator that aggregates dominance degrees. Its calculation, as shown in Equation 7, involves a position-based weighting.

$$F_Q(a_1, \dots, a_n) = \sum_{i=1}^{n} w_i b_i \quad (7)$$

Ultimately, the alternative with the highest total dominance score will be designated as the "Expert Priority" for that specific symptom item.

The final stage in the computational model is to integrate the expert priority results with the patient's self-report data. This mechanism is designed based on the principle of a context-aware system that responds to triggers dynamically. Specifically, this approach adopts the Event-Condition-Action (ECA) model, where the analysis of an item will only be executed if the "condition" is met, i.e., the patient's score is greater than zero. This conditional activation ensures that the system only focuses on symptoms that are relevant and experienced by the patient [23]. A systematic review of the use of Patient-Reported Outcome Measures (PROMs) like DASS has shown that the most effective interventions are those that use PROMs to monitor disease symptoms and directly link them to care pathways, a mechanism applied in this model [24]. Subsequently, the raw score reported by the patient (ranging from 0-3) is not used directly. Instead, the score is normalized to create a "severity" value by dividing it by the maximum possible score (3). This resulting severity weight (e.g., a score of '2' becomes 2/3 = 0.66), which ranges from 0 to 1, is then used to weight the final expert priority vector. While the 0-3 numerical scale is common in clinical practice, this normalization process ensures the patient's input acts as a proportional weight (from 0% to 100%) against the established expert consensus. The use of a numerical scale to measure symptom severity has precedent in clinical practice, such as in the Fibromyalgia diagnosis criteria which use a Symptom Severity (SS) scale with a 0-3 range to assess the severity level of key symptoms [25]. Mathematically, this dynamic synthesis process is described in Equation 8.

$$V_k = \sum_{i=1}^{21} S_i \cdot \mathbb{I}(k_i^* = k), \quad (8)$$

where $k_i^* = \arg\max_{k' \in K} (Q_i(k'))$

The combination of conditional activation based on ECA rules and severity-based weighting ensures a more personal, efficient, and relevant assessment for each individual.

$$FinalScore_k = \frac{V_k}{\sum V_k} \quad (9)$$

After the final score for each alternative (Depression, Anxiety, Stress) is accumulated in the crisp vector V_k (result of Equation 8), the last step is to perform normalization. This process aims to convert the raw accumulated scores into a proportional distribution (e.g., in percentage form) that can be compared directly. The alternative with the highest $FinalScore_k$ value (result of Equation 9) will be determined as the final diagnostic interpretation from the model. This result indicates the most dominant condition (Depression, Anxiety, or Stress) for that patient, based on the established expert consensus and the patient's self-report data.

2.4 Implementation

The model implementation process was carried out by translating the methodology flow into functional program scripts using the Python programming language. The implementation stage begins by executing the function to calculate the capability scores and importance weights (IOWA Weight) for each expert based on the prepared data.

Subsequently, the system runs an iterative process for each of the 21 DASS items. In each iteration, functions are called sequentially: the function to transform expert

preferences into a fuzzy matrix, followed by the group aggregation function to produce a collective preference matrix, also in the form of a TFN matrix. This matrix is then processed by the QGDD ranking determination function to produce the "Expert Priority". Finally, conditional logic is implemented where the patient's self-report score is used to accumulate the final score for the alternative that has been prioritized by the model.

3. Results and Discussions

This section presents the implementation results of the computational model developed in the research methodology. The analysis focuses on each stage of the system, from the objective weighting of the five clinical experts to the final diagnostic interpretation of respondent data. The results demonstrate how the integration of SAW and Fuzzy-IOWA reduces subjectivity in DASS-21 score interpretation.

3.1 Final Expert Weight Calculation Results

This subsection details the implementation of Phase 2 of the research framework (Figure 1): Expert Capability Quantification and Weight Determination. This process converts raw capability data from five experts (Table 1) into standardized Final Weights (w_i), representing each expert's "voting power" or influence within the group decision model.

The initial phase of this analytical procedure involves normalizing the raw expert capability data by applying the benefit-criteria equation derived from the Simple Additive Weighting (SAW) method. This normalization process is fundamentally critical to ensure that all evaluative criteria which possess highly varied numerical ranges, spanning from 'Patient Volume' in the hundreds to 'Education Level' in single digits are consistently mapped into a uniform range between 0 and 1. By executing this dimension-free transformation, the unit bias inherent in the differing measurement scales across criteria is effectively eliminated, thereby facilitating an objective and equitable cross-criteria comparison during the subsequent data integration stages. The systematic results of this detailed normalization for each expert capability parameter are presented for further verification in Table 2.

Table 2. Normalized Expert Capability Matrix r_{ij}

Psychologist	Years of Experience	Patient	Education	Publication
P1	1.0000	0.0000	1.0000	1.0000
P2	0.6667	1.0000	1.0000	0.0323
P3	0.0000	0.2268	1.0000	0.0000
P4	0.1667	0.4845	1.0000	0.0968
P5	0.3333	0.3986	1.0000	0.0323

As shown in Table 2, P1 achieves the maximum score in 'Experience' and 'Publications', while P2 dominates the 'Patients' criterion. Notably, all five experts possess equivalent 'Education' levels (S2), resulting in a uniform normalized value of 1.0 for that criterion.

After obtaining the normalized matrix (r_{ij}) in Table 2, the initial capability score (V_i) for each expert is calculated using Equation (2). This calculation is a weighted sum, where each normalized criterion value is multiplied by its respective criterion Experience (0.38), Patients (0.25), Education (0.25), and Publication (0.12). The calculation example for Expert 1 (P1) is:

$$V_1 = (1.0000 \times 0.3800) + (0.0000 \times 0.2500) + (1.0000 \times 0.2500) + (1.0000 \times 0.1200) = 0.7500$$

The complete initial score calculation results for all five experts are presented in Table 3. The results in Table 3 indicate that P2 (0.7582) and P1 (0.7500) possess the highest and nearly identical initial capability scores, while P3 (0.3075) has the lowest. These initial scores are fundamentally determined by the criteria weights assigned to each parameter. In this study, the criteria weights were not arbitrarily assigned but were derived from the average percentage of influence provided by all five experts in the group. By averaging their professional judgment, the importance of each criterion was established as: Experience (38%), Patients (25%), Education (25%), and Publication (12%). This collective weighting ensures that the model reflects the group's consensus on what constitutes expert authority in clinical psychology.

Table 3. Initial Expert Capability (V_i)

Psychologist	Calculation (Criteria Weight × Value)	Initial Score (V_i)
P1	(0.38*1.00) + (0.25*0.00) + (0.25*1.00) + (0.12*1.00)	0.7500
P2	(0.38*0.67) + (0.25*1.00) + (0.25*1.00) + (0.12*0.03)	0.7582
P3	(0.38*0.00) + (0.25*0.23) + (0.25*1.00) + (0.12*0.00)	0.3075
P4	(0.38*0.17) + (0.25*0.48) + (0.25*1.00) + (0.12*0.10)	0.4466
P5	(0.38*0.33) + (0.25*0.40) + (0.25*1.00) + (0.12*0.03)	0.4790
Total (V_i)		2.7413

The final stage is to convert these initial capability scores (V_i) into proportional final weights (w_i) that sum to 1.0. This transformation is conducted using Equation 3where each expert's score is divided by the total sum of all scores (2.7413). This normalization step ensures that each expert's influence is represented as a relative probability or "voting power." Table 4 presents the final expert weight calculation results, representing the definitive influence of each expert in the subsequent aggregation process.

The results in Table 4 represent the final distribution of influence or "voting power" assigned to each expert. By transforming the initial capability scores into proportional weights, the system ensures that the group decision-making process is balanced and scientifically accountable.

Based on this distribution, visualized in Figure 2, it is evident that P2 (27.66%) and P1 (27.36%) hold the most significant influence in the model, collectively

accounting for over 55% of the total voting power. Meanwhile, P3 holds the lowest influence at 11.22%. These final weights (w_i) serve as the critical "inducing variables" for the next phase: the Fuzzy-IOWA Aggregation, where these weights will dictate how each expert's preference is prioritized in forming the group consensus.

Table 4. Final Expert Weight Calculation Results (w_i)

Psychologist	Initial Score(V_i)	calculation (V_i/ V_t)	Final Weight
P1	0.7500	0.7500 / 2.7413	0.2736
P2	0.7582	0.7582 / 2.7413	0.2766
P3	0.3075	0.3075 / 2.7413	0.1122
P4	0.4466	0.4466 / 2.7413	0.1629
P5	0.4790	0.4790 / 2.7413	0.1747
Total (V_t)	2.7413		1.0000

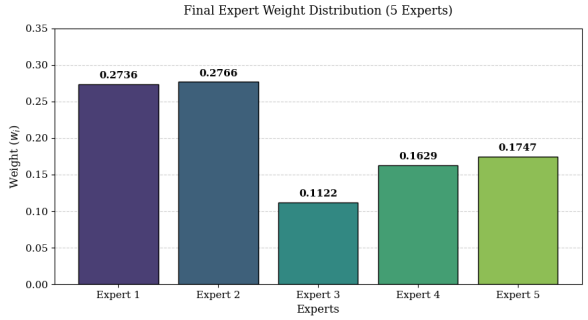


Figure 2. Final Expert Weight Distribution

3.2 Expert Consensus Model Results

After determining the final expert weights in Table 4, the model proceeds to the Fuzzy-IOWA Aggregation Phase, where the weights (w_i) act as inducing variables to aggregate the preferences of the five experts. This process utilizes the Fuzzy-IOWA operator, as defined in Equation 4, for each of the 21 DASS-21 items. The output is a collection of 21 priority vectors, each consisting of a 1×3 matrix for Depression, Anxiety, and Stress, which collectively form the Final Expert Consensus Model. This model serves as the standardized weighting matrix or answer key used to interpret respondent scores in the subsequent stage.

A visualization of the 21 aggregated priority vectors is presented in the heatmap shown in Figure 3. A heatmap was selected as a visualization tool due to its ability to provide a comprehensive overview of complex data, where a third quantitative dimension, specifically the priority score, is visually captured through color variations. This visualization captures the unique characteristics of each DASS-21 item as defined by the professional panel. A lighter color intensity, which approaches a yellow or white hue, indicates a higher priority score closer to 1.0. For example, in Item 3, relating to the inability to experience positive feelings, the heatmap displays a very light color for the Depression dimension. This confirms a strong expert consensus that this item is a primary indicator for depression. Conversely, for Item 1, regarding the difficulty in winding down, the priority is distributed between Anxiety and Stress, reflecting the clinical comorbidity recognized by the experts.

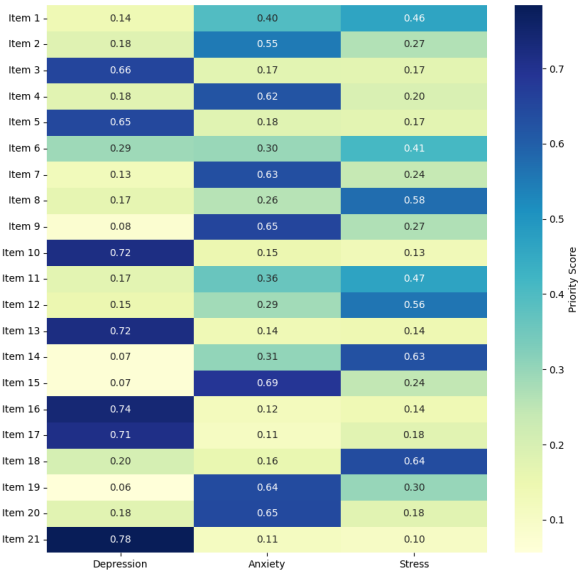


Figure 3. Heatmap of Priority Aggregation Vector per Item (Final Consensus Model)

To analyze the overall tendency of the group consensus, the Global Dominance Degree (QGDD) method was applied to the 21 aggregation vectors. The results in Figure 4 reveal that the Final Expert Consensus Model has a dominant overall tendency toward Anxiety, with a global score of 0.3566, followed by Depression at 0.3346 and Stress at 0.3089. This dominance of Anxiety can be further explained by both methodological and domain-specific factors. From a methodological perspective, the Fuzzy-IOWA operator assigns greater influence to experts with higher capability scores, particularly Expert 1 and Expert 2. If these dominant experts exhibit a tendency to associate ambiguous symptoms with anxiety-related interpretations, this preference will be amplified during aggregation.

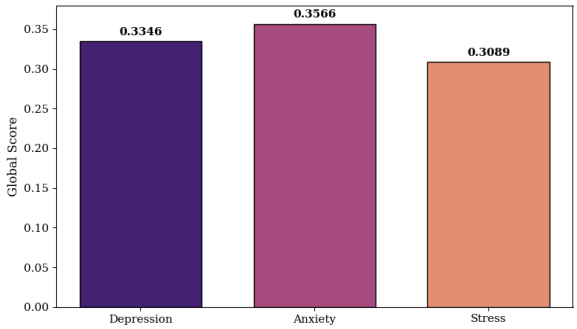


Figure 4. Global QGDD Alternative Scores (Consensus Model Characteristics)

From a domain perspective, Anxiety indicators within the DASS-21 framework are often more explicitly linked to physiological arousal symptoms, such as restlessness and tension, which tend to be more consistently recognized across experts. In contrast, depression and stress may involve broader and more overlapping symptom definitions, making them less dominant in aggregated preference structures.

Therefore, the observed dominance is not solely a result of the aggregation mechanism, but rather an interaction between expert weighting and the inherent characteristics of the DASS-21 constructs.

This tendency aligns with the physiological hyperarousal component of the Tripartite Model, where symptoms of autonomic arousal are more distinctly mapped to anxiety by clinicians compared to the generalized symptoms of stress. By prioritizing the perspectives of practitioners with the most extensive clinical experience, specifically Expert 1 and Expert 2, through the IOWA mechanism, the model successfully captures this nuanced professional preference.

However, it is important to note that the dominance pattern and aggregated results are also influenced by the relatively small number of experts and the structured nature of the DASS-21 instrument. In scenarios where expert opinions are relatively homogeneous, the final aggregated outputs may appear similar to those produced by simpler methods such as OWA or arithmetic averaging.

The practical advantage of the proposed Fuzzy-IOWA approach becomes more evident in situations where there is greater variability in expert expertise or disagreement in interpretation. In such cases, the ability to incorporate differentiated expert influence provides a more structured and theoretically grounded aggregation mechanism compared to equal-weight approaches.

3.3 Patient Score Integration Results

This is the final stage of the research framework (Figure 1), where the constructed GDSS model is applied to DASS-21 respondent data to generate a diagnostic interpretation. This process integrates two distinct data sources: the Final Expert Consensus Model from Figure 3, which is a 21x3 (D, A, S) priority matrix reflecting the collective professional weighting of the five experts, and the respondent's raw scores, which consist of the patient's individual responses valued at 0, 1, 2, or 3 for each of the 21 DASS-21 items.

The model performs the integration process using a weighted aggregation method. For each item ($i = 1$ to 21), the raw item score (x_i) is normalized into a severity index by dividing it by the maximum possible score (3). This severity index is then multiplied by the corresponding priority vector from the Expert Consensus Model. The resulting values are accumulated into total scores for Depression (D_{total}), Anxiety (A_{total}), and Stress (S_{total}).

As an illustrative case, we analyze the results for Respondent 1. After processing all 21 items, the accumulated raw scores for Respondent 1 are visualized in Figure 5. Figure 5 displays the raw accumulated weights for each domain before final normalization. To provide a clear and intuitive diagnostic interpretation, these raw scores are converted into a percentage distribution using a normalization function. This final step ensures that the diagnostic output is represented as a probability distribution where D + A + S = 100%. The final interpretation results for Respondent 1 are presented in Figure 6.

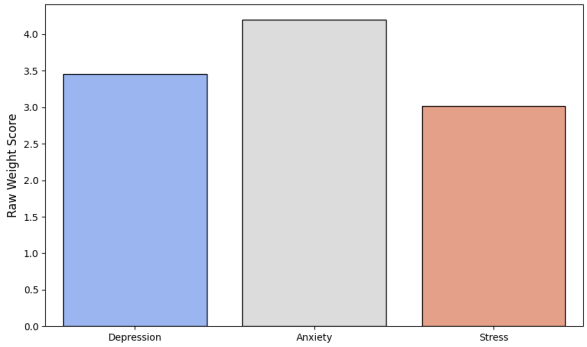


Figure 5. Respondent 1 Raw Accumulated Score Chart (Before Normalization)

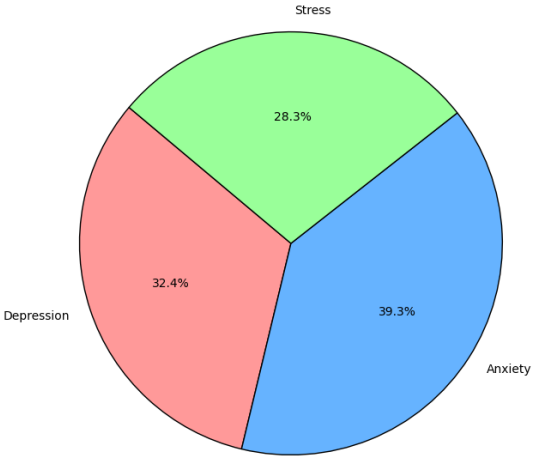


Figure 6. Final patient score distribution chart (case study 'respondent 1')

Table 5 illustrates the model's diagnostic interpretation across different cases. It is apparent that Anxiety was consistently identified as the dominant condition for all analyzed profiles. However, the probability distribution between Depression, Anxiety, and Stress varies significantly for each individual. For example, the gap between Anxiety and Depression is more pronounced in Respondent 1 compared to Respondent 3.

Table 5. Summary of Final Patient Alternative Scores (GDSS Model Normalization Results)

Respondent	Depression	Anxiety	Stress
Respondent 1	0.3152	0.5094	0.1754
Respondent 2	0.3341	0.4425	0.2234
Respondent 3	0.3418	0.4382	0.2200
Respondent 4	0.3482	0.4451	0.2067

This variation demonstrates that the GDSS model successfully produces a unique and specific interpretation profile for each respondent. By integrating the self-reported severity from the patient with the high-level professional consensus of the five experts, the system provides a structured aggregation of expert interpretations that reflects established diagnostic reasoning patterns. This approach effectively addresses the ambiguity often found in traditional

DASS-21 assessments where symptoms may overlap across different psychological dimensions.

3.4 Comparative Evaluation of Aggregation Logic

To analyze the operational characteristics of the proposed model, this section provides a two-layer analysis: a macro-level output comparison (Figure 7) and a micro-level analysis of weighting behavior (Figure 8). This comparison involves three distinct approaches: the proposed Fuzzy-IOWA (utilizing dynamic expert weighting), Standard OWA (utilizing uniform expert weighting), and the Standard Average (arithmetic mean).

Figure 7 illustrates the diagnostic probabilities for 20 respondents across the three tested methods. At first glance, the results show a high degree of numerical convergence. For instance, in Respondent 1, the Anxiety probability yields 0.5094 for IOWA, 0.4912 for OWA, and 0.4850 for the Standard Average. This phenomenon suggests that while the aggregation methods differ, the final accumulated weights for a group of five experts tend to smooth out extreme preferences across 21 items. This convergence can be explained by the law of aggregation in multi-expert systems, where increasing the number of items and contributors tends to reduce variance in the final output. In this study, the use of 21 DASS items combined with five expert inputs results in a smoothing effect, where extreme preferences are averaged out regardless of the aggregation method.

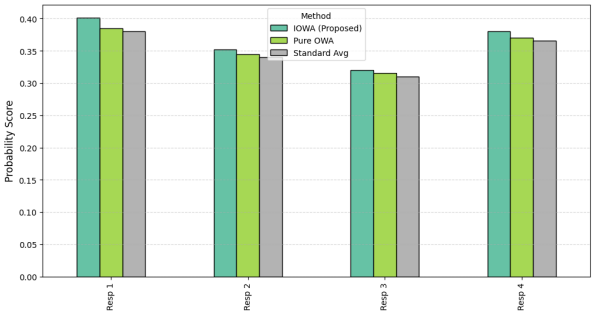


Figure 7. Comparative Diagnostic Probabilities for Anxiety Domain

As a result, differences between Fuzzy-IOWA, OWA, and arithmetic averaging become less pronounced at the final output level as shown in Figure 8. This highlights that output similarity does not necessarily indicate methodological equivalence, but rather reflects the stabilizing effect of repeated aggregation across multiple dimensions.

From a practical perspective, this convergence indicates that in structured assessment scenarios with relatively homogeneous expert profiles, simpler aggregation methods may produce results that are numerically similar to more advanced approaches. However, the advantage of the proposed Fuzzy-IOWA approach becomes more evident in situations where there is significant variation in expert expertise or disagreement in interpretation. In such cases, equal-weight aggregation methods may fail to adequately represent the influence of more experienced practitioners. Therefore, the contribution of this model should be understood as an enhancement in aggregation structure and interpretability, rather than a guaranteed improvement in numerical outcomes.

To examine the internal mechanics of the model, Figure 8 plots the contribution of expert opinions for each DASS-21 item. The significant fluctuations in the IOWA line compared to the relatively flat Standard Average line demonstrate the active operation of the Dynamic Inducing Mechanism. In this framework, the consensus is systematically influenced by practitioners with higher objective credibility scores, such as Expert 1 and Expert 2.

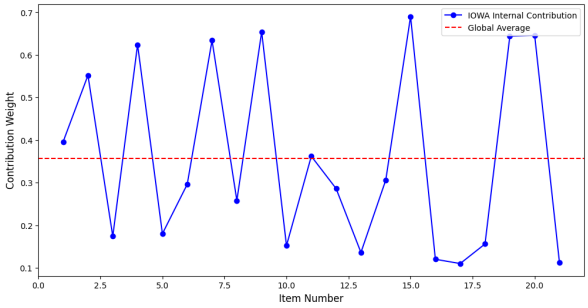


Figure 8. Logical Aggregation Analysis: Internal Logic Weight over 21 Items

Unlike standard arithmetic averaging, which treats all inputs equally, the IOWA operator prioritizes interpretations from experts with more extensive clinical records and qualifications during cases of item ambiguity. While this does not yet constitute clinical validation against external gold standards, the analysis suggests that the Fuzzy-IOWA model provides an aggregation that mirrors a hierarchical consultation structure. This allows the influence of each expert within the group to be mathematically aligned with their quantified professional background.

4. Conclusions

This study presents a Group Decision Support System (GDSS) framework for interpreting DASS-21 scores by integrating expert-based weighting and Fuzzy-IOWA aggregation. The proposed approach introduces a structured mechanism for incorporating expert capability into the decision-making process, where the influence of each professional is quantified using the SAW method; for instance, Expert 2 (27.66%) and Expert 1 (27.36%) were identified as having the highest influence within the group. The model generates interpretable distributions, such as the Anxiety score of 0.5094 for Respondent 1, reflecting a consistent consensus structure where experts with higher quantified capability contribute more significantly to the final interpretation. While these results show numerical similarity with conventional averaging, further analysis reveals critical differences in the internal aggregation logic.

The proposed framework contributes to the development of decision support models by providing a transparent and systematic approach for handling subjective expert judgments, particularly in cases involving ambiguity and overlapping symptoms. By providing a mathematically accountable process, the system bridges the gap between subjective clinical intuition and objective computational modeling. However, it is important to emphasize that this study focuses primarily on the modeling and aggregation mechanism rather than formal clinical validation against external gold standards. As such, the system is designed to enhance the transparency of group consensus building rather than to replace existing diagnostic procedures.

Several limitations must be acknowledged. The model has not yet been validated against external clinical diagnoses or longitudinal patient outcomes, and the relatively small number of experts and respondents, combined with capability criteria that may not capture all dimensions of clinical expertise, may limit the generalizability of the findings. Additionally, the system was not benchmarked against data-driven approaches such as machine learning or deep learning classifiers, which have demonstrated competitive performance in mental health assessment tasks. Therefore, the proposed system should be interpreted as a methodological framework for structured expert aggregation rather than a clinically validated diagnostic tool. Future research should address these limitations by: (1) conducting external validation using real clinical data paired with professional diagnoses; (2) expanding the study with larger expert panels and more diverse respondent datasets to enhance generalizability; and (3) incorporating comparative studies with machine learning or deep learning approaches to more rigorously position the GDSS framework relative to current state-of-the-art methods. Addressing these directions will strengthen the scientific credibility of the proposed model and better define its role as a complementary tool within clinical decision-making workflows.

Acknowledgements

This research was funded by Universitas Sebelas Maret, Surakarta Indonesia, managed under The Applied Excellence Research Grant program, under Contract No. 369/UN27.22/PT.01.03/2025.

References

[1] P. F. Lovibond and S. H. Lovibond, “The Structure of Negative Emotional States: Comparison of The Depression Anxiety Stress Scales (DASS) With The Beck Depression and Anxiety Inventories,” *Behaviour Research and Therapy*, vol. 33, no. 3, pp. 335-343, 1995, doi: 10.1016/0005-7967(94)00075-U

[2] E. Wahyuni and W. Cahyawulan, “Factorial Structure of The Depression, Anxiety, Stress Scale (DASS) Among University Students: The Impact of Counseling Services,” *in Proc. 8th Int. Conf. Educ. Innov. (ICEI)*, 2025, pp. 460–474. doi: 10.2991/978-2-38476-360-3_44.

[3] I. P. Clara, B. J. Cox, and M. W. Enns, “Confirmatory Factor Analysis of the Depression-Anxiety-Stress Scales in Depressed and Anxious Patients,” *Journal of Psychopathology and Behavioral Assessment*, vol. 23, no. 1, pp. 61–67, 2001, doi: 10.1023/A:1011095624717.

[4] J. Zhan, K. Zhang, P. Liu, and W. Pedrycz, “A novel group decision-making approach in multi-scale environments,” *Applied Intelligence*, vol. 53, no. 12, pp. 15127–15146, Jun. 2023, doi: 10.1007/s10489-022-04279-5.

[5] A. Priya, S. Garg, and N. P. Tigga, “Predicting Anxiety, Depression and Stress in Modern Life using Machine Learning Algorithms,” *Procedia Computer Science*, vol. 167, pp. 1258–1267, 2020, doi: 10.1016/j.procs.2020.03.442.

[6] S. R. Marsidi, “Identification of Stress, Anxiety, And Depression Levels of Students In Preparation for The Exit Exam Competency Test,” *Journal of Vocational Health Studies*, vol. 5, no. 2, pp. 87-93, Nov. 2021, doi: 10.20473/jvhs.v5.i2.2021.87-93.

[7] Shikha Verma, Aman Goyal, and Santushti Gandhi, “Predicting Stress, Anxiety and Depression Among the University Students of India Post-Covid,” *Int J Innov Sci Res Technol*, vol. 8, no. 8, pp. 2244-2250, 2023, doi: 10.5281/zenodo.8330745

[8] R. T. Sutton, D. Pincock, D. C. Baumgart, D. C. Sadowski, R. N. Fedorak, and K. I. Kroeker, “An overview of clinical decision support systems: benefits, risks, and strategies for success,” *npj Digital Medicine*. vol. 3, no. pp. 1-10, 2020, doi: 10.1038/s41746-020-0221-y.

[9] M. Salimparsa, K. Sedig, D. J. Lizotte, S. S. Abdullah, N. Chalabianloo, and F. T. Muanda, "Explainable AI for clinical decision support systems: Literature review, key gaps, and research synthesis," *Informatics*, vol. 12, no. 119, pp. 1–27, 2025, doi:10.3390/informatics12040119.

[10] L. Ali, M. S. Sadi, and O. Goni, “Diagnosis of heart diseases: A fuzzy-logic-based approach,” *PLoS One*, vol. 19, no. 2, pp. 1-25, Feb. 2024, doi: 10.1371/journal.pone.0293112.

[11] R. R. YAGER, “On Ordered Weighted Averaging Aggregation Operators in Multicriteria Decisionmaking,” in Readings in Fuzzy Sets for Intelligent Systems, D. Dubois, H. Prade, and R. R. Yager, Eds. San Mateo, CA, USA: Morgan Kaufmann, 1993, pp. 80–87, doi: 10.1016/B978-1-4832-1450-4.50011-0.

[12] R. R. Yager, and D. P. Filev, “Induced Ordered Weighted Averaging Operators,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 29, no. 2, pp. 141-150, April 1999, doi: 10.1109/3477.752789.

[13] F. Chiclana, E. Herrera-Viedma, F. Herrera, and S. Alonso, “Some induced ordered weighted averaging operators and their use for solving group decision-making problems based on fuzzy preference relations,” *Eur J Oper Res*, vol. 182, no. 1, pp. 383–399, Oct. 2007, doi: 10.1016/j.ejor.2006.08.032.

[14] X. Tong and Z. J. Wang, “A group decision framework with intuitionistic preference relations and its application to low carbon supplier selection,” *Int J Environ Res Public Health*, vol. 13, no. 9, pp. 1-16, Sep. 2016, doi: 10.3390/ijerph13090923.

[15] Wiharto *et al.*, “Model Group Decision Support System Based on Depression Anxiety Stress Scales Using Ordered Weighted Averaging Aggregation Method,” *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, vol. 7, no. 2, pp. 307–317, Apr. 2025, doi: 10.35882/jeeemi.v7i2.678.

[16] E. Kipoulas, I. Edwards, R. Radakovic, and P. I. Beazley, “Perceptions of bias and credibility of male and female clinical psychologist and psychiatrist expert witnesses presenting clinical information in the courtroom,” *Int J Law Psychiatry*, vol. 96, pp. 1-14, Sep. 2024, doi: 10.1016/j.ijlp.2024.102016.

[17] W. Lu, C. Liang, and Y. Ding, *A Method Determining the Objective Weights of Experts Based on Evidence Similarity in Group Decision-making*,” in Proc. 4th Int. Conf. Wireless Commun., Netw. Mobile Comput., 2008, doi: 10.1109/WiCom.2008.2764.

[18] A. Peña, “The Dreyfus model of clinical problem-solving skills acquisition: a critical perspective.,” *Medical Education Online*, 2010. doi: 10.3402/meo.v15i0.4846.

[19] M. Younan and K. A. Martire, “Likeability and Expert Persuasion: Dislikeability Reduces the Perceived Persuasiveness of Expert Evidence,” *Front Psychol*, vol. 12-2021, Dec. 2021, doi: 10.3389/fpsyg.2021.785677.

[20] R. Brammer, “Expertise in Psychological Inquiry: A Computerized Assessment,” ERIC Document ED410261, Mar. 1997. [Online]. Available: https://eric.ed.gov/?id=ED410261 .

[21] M. I. Panjaitan, "Simple Additive Weighting (SAW) method in Determining Beneficiaries of Foundation Benefits," Login: Jurnal Teknologi Komputer, vol. 13, no. 1, pp. 19–25, Jun. 2019. [Online]. Available: https://media.neliti.com/media/publications/326766-simple-additive-weighting-saw-method-in-f8f093e8.pdf

[22] F. Herrera and E. Herrera-Viedma, "Linguistic decision analysis: steps for solving decision problems under linguistic information," *Fuzzy Sets Syst.*, vol. 115, no. 1, pp. 67–82, 2000, doi: 10.1016/S0165-0114(99)00024-X.

[23] N. Hussain, H. H. Wang, and C. Buckingham, "Policy based generic autonomic adapter for a context-aware social-collaborative system," in *2018 International Conference on Intelligent Systems and Computer Vision (ISCV), Fez, Morocco,* 2018, pp. 1-9, doi: 10.1109/ISACV.2018.8354044.

[24] J. M. Bonsel, A. J. Itiola, A. S. Huberts, G. J. Bonsel, and H. Penton, "The use of patient-reported outcome measures to improve patient-related outcomes - a systematic review," *Health Qual Life Outcomes*, vol. 22, no. 1, pp. 1-19 Dec. 2024, doi: 10.1186/s12955-024-02312-4.

[25] S. Al Sharie *et al.*, "Unraveling the Complex Web of Fibromyalgia: A Narrative Review," *Medicina*, vol. 60, no. 2, Art. no. 272, 2024, doi: 10.3390/medicina60020272.