Cracking the Code: Examining Psychometric Rigor of the Provider Decision Process Assessment Instrument (PDPAI) among Residents' Trainees and Expert Physicians

Sana Younas*, Saeeda Khanum**

Department of Behavioral Sciences,
School of Social Sciences and Humanities,
National University of Sciences and Technology, Islamabad Pakistan.
*e-mail: sanayounasmphil@gmail.com.
**e-mail: saeeda.khanum@s3h.nust.edu.pk.

Submitted: 29.03.2024 - **Accepted:** 29.07.2024

Available online: 28.11.2024

Abstract

The realm of healthcare decision-making remains inadequately explored, specifically in assessing the psychometric characteristics of tools like the Provider Decision Process Assessment Instrument (PDPAI). This study aims to fill this void by examining decisional conflict among resident trainees and experienced consultant physicians. We approached a total of 347 physicians using a convenient sampling method from tertiary care hospitals. The analysis encompassed (i) factorial validity of PDPAI through confirmatory factor analysis (CFA) and evaluating the single group CFA models and (ii) multigroup CFA models, (iii) examining factorial invariance among residents' trainees and experienced physicians' groups, (iv) Rasch analysis assessing the individual item impact on the subdomains, (v) internal consistency (vi) convergent and discriminant validity. The bi-factor model adequately fit the data as all factor loadings (0.44-0.70) were statistically significant ($p < 0.05$). The bifactor model supported the global construct or the sub-domains as suitable measurement models. The PDPAI showed invariance for use across two physician groups. Physicians encountered greatest difficulty in item "I was clear what treatment would be best for this patient." [MNSQ Infit/Outfit: 1.327/1.278] and found the easiest item "It was easy to identify all of the considerations that affect the decision" [0.902/0.869]. Adequate internal consistency was revealed through

ISSNe 1972-5620, Doi: 10.3280/rip2024oa18309

Cronbach and Omega coefficient values. Convergent and discriminant validity of PDPAI was supported by correlating with team decision making questionnaire and compassion fatigue respectively. The PDPAI's validated cross-group invariance highlights it's applicability to a diverse range of physician groups, guiding tailored interventions.

Keywords: healthcare physicians, factorial validity, Rasch analysis, Internal consistency, convergent and discriminant validity

Introduction

This study endeavors to fill a crucial void in existing literature by scrutinizing the psychometric attributes of the 12-item Provider Decision Process Assessment Instrument (PDPAI) (Dolan, 1999) within the cohorts of resident trainees and experienced consultant physicians. The identified critical gap pertains to the insufficient understanding of decisional conflict among these groups – conflict stemming from uncertainty surrounding potential action plans, as articulated by Han et al., (2019). Despite the existence of a number of decision evaluation scales (for e.g., PDPAI (Dolan, 1999), Decision Attitude Scale (Barry et al., 1997), the Satisfaction with decision making process questionnaire (Sainfort & Booske, 2000) etc.), a thorough examination into their robustness, psychometric properties, and application in various health settings and populations remains conspicuously absent. There is a lack of research that requires an investigation, into verifying decision assessment scales especially in healthcare settings to ensure their accuracy and consistency. This gap is highlighted by the number of studies focusing on the characteristics of PDPAI, among both early career and experienced physicians, also known as consultants or attending physicians. These professionals offer guidance, supervision, and specialized knowledge to trainees during their training period (Younas et al., 2023: Younas & Khanum, 2024a, b). By exploring the intricacies of decision making in these circles, we may gain insights into reducing decision related conflicts and ultimately improving patient outcomes in healthcare settings.

Decisional conflict results when a person is uncertain about a potential course of action (Moure et al., 2023; O'Connor et al., 2002). Additionally, decisional conflict is aggravated when people compromise on values in order to choose a course of action or when they anticipate feeling re-morseful for not selecting certain options (Liu et al., 2023). The main behavioral manifestations of decisional conflict are expressed uncertainty about possibilities, expression of unfavorable consequences of alternatives, and a propensity for delaying decisions (Dhami & Mandel, 2022). Decisional conflict is contributed by factors such as lack of support or external pressure, unclear personal values, and limited knowledge. Decisional conflict can be effectively reduced by decision supporting interventions. People feel more informed when they are given information about options, and side effects (Liao et al., 2023). Detailed explanations of outcomes, including their effects on the body, mind, and spirit, are a key component of value-clarification strategies. People are also urged to evaluate the significance of these outcomes for themselves. Guidance during the stages of collaborative decision-making promotes support in decision-making. As a result, uncertainty resulting from these adaptive aspects decreases, giving the impression that the option was better (Légaré et al., 2006; O'Donnell et al., 2023). This improved choice is distinguished by a sense of better knowledge, agreement with personal beliefs, increased likelihood of adherence, and higher levels of satisfaction. The beneficial effects of decision-supporting treatments on decisional conflict and the associated changeable dimensions are strongly supported by empirical data (O'Connor, 1995; Wendler & Rid, 2011).

By addressing the presumptive causes of this conflict, decision aids have the potential to reduce decisional conflict (O'Conner, 2010). For instance, a thorough examination of the options and a thorough investigation of the prospective outcomes can help to offset the lack of information. Physicians'confidence in the informed nature, alignment with personal values, and feasibility of their decisions is gauged through the efficacy of their decision making. The existing scale evaluates how comfortable physicians feel about their decisions. The importance of examining medical decision making processes in healthcare environments is emphasized due to the potential influence on patient results and healthcare procedures.

The Health Physician's Version of the Decisional Conflict Scale (DCS)

Dolan (1999) first adapted the DCS among 22 healthcare workers (Trees et al., 2017). Numerous studies emphasized the relevance of decisional conflict as a marker of high-quality decision-making among medical professionals. Accordingly, the adapted version of the DCS, called the Provider Decision Process Assessment Instrument (PDPAI), was developed to measure decisional conflict among medical professionals. According to Zimmer-Watson et al. (2008), PDPAI fills a significant research gap, by gauging individual medical professionals' perceptions of the decision process. The adapted instrument comprised four more items, in addition to the original 16, resulting in a 20-item instrument where eight of the 20 items were the same as in DCS (i.e., items 9, 10, 13, 18, 19, 21, 22, and 25). It showed an acceptable construct validity and reliability, as revealed through the Cronbach's alpha of 0.90. Another study (Honmg et al., 2003) translated the PDPAI into French and administered it to 34 family physicians in Quebec, resulting in a Cronbach's alpha coefficient of 0.82. Furthermore, nine out of the original 16 items of DCS were used in an Australian study that investigated the ability of family physicians to help patients decide about prostate-specific antigen screening (as cited in Pecanac et al., 2018). This study also showed consistency of the instrument as reveled through the Cronbach's alpha value ($\alpha = 0.81$).

However, there are some shortcomings of the scale. First, Dolan (1999) recruited only 112 participants resulting in limited generalizability and implications of the sample. Second, the author did not explore the factor structure of the scale to establish the construct validity rather used Spearman correlation and found negative correlation between decision conflict and satisfaction. Third, Dolan's study has been cited in 67 research articles, however, to the best of our knowledge, none of the studies worked on factor structure of the scale. Some of the review studies, however, suggested that the PDPAI instrument exhibits good internal consistency in terms of psychometric quality, but also suggested that validity has not been sufficiently examined (Scholl et al., 2011; Simon, Loh, & Härter, 2007). Hence, it was necessary to validate the scale in an Asian culture before using it.

Cultural traditions in collectivistic nations such as Pakistan, emphasize group decision-making that frequently affect medical decisions. For instance, decisions about how to manage elderly patients may involve extensive family influence, influencing both end-of-life and medical care options (García & Garasic, 2021). The PDPAI has been validated to ensure that it appropriately captures the subtleties of decision difficulty encountered by physicians while respecting the cultural setting in which they work. Physicians may encounter potential challenges influencing their ability to make well-informed decisions such as a lack of access to literature or information gaps. The validation of this scale assures that the evaluation properly captures the ethical issues inherent in physicians'decision-making processes.

The objective of this investigation was to evaluate the psychometric characteristics of the 12-item PDPAI through the application of both classical test theory and item-response theory. We examined the congruence of the domains using confirmatory factor analysis for (i) single factor PDPAI (ii) two factor PDPAI and (iii) bi-factor PDPAI. We also examined the PDPAI for measure equivalency in early career and experienced physicians. Since, it has been emphasized that years of experience in uncertain medical situations helps one to make better decisions (Falzer, 2018; Klein, 2015). Individuals having more experience use more intuitive thinking, cues, and recognize the situation through pattern formation as compared to the young physicians who have less experience of dealing with uncertain situations in medical settings (Epstein, 2011; Ruzsa et al., 2020). Young trainee physicians having less experience may get more stressed in uncertain and time pressured situations, thus having more conflict with their decisions as compared to experienced physicians (for instance, consultants) who are less stressed and are more satisfied with their decisions. Additionally, Rasch modeling was used alongside confirmatory factor analysis. Wilson et al. (2006) noted, Rasch modeling provides a direct assessment of the relationship between respondents' positions and the placement of items on the latent variable scale. Contrarily, confirmatory factor analysis, as outlined by Strauss and Smith (2009), examines the connections between components. Omega reliability, convergent and discriminant validity of the PDPAI were also assessed.

Methods

Sample

Ethical approval was obtained prior to data collection from ethical review board of the University. The research was a descriptive cross-sectional survey. Data collection started in March 2022 and ended till January 2023 from all major tertiary care hospitals in the Potohar region located northeast of Pakistan. Convenient sampling technique was used to approach the participants. Greater sample size is considered better for validation of a measure. Hence, the 10:1 minimum standard (10 cases per parameter) was used to draw the sample size (Hair et al., 2010). Total 347 practitioners participated in the study. Among them, (n = 180) were residents' trainees who were referred as early career physicians while (n = 167) were senior and experienced physicians including classified consultants (Assistant Professors, Associate Professors, and Professors). Their age ranged from 24-66 years (M = 32.066; SD = 7.513). Resident trainees had experience between 1 and 4 years while senior physicians have 6 and more years of experience.

Measures

Provider Decision Process Assessment Instrument (PDPAI) developed by Dolan (1999) to measure the healthcare provider assessment of decision-making. The instrument is in English, hence did not require translation as Pakistani physicians can understand and speak English well. The assessment comprises a total of 12 items, with scoring conducted on a five-point Likert scale. It explores the perceived complexity present in situations involving decision-making and acknowledges the intricate interactions between variables that may make medical decisions more challenging.

Additionally, Dolan's scale evaluates clinicians'knowledge of patients'preferences and values during the decision-making process, further emphasizes the significance of patient-centered care. It examines how much medical practitioners value shared decision-making, respects patients'autonomy, and involves patients in treatment conversations. Some of the examples of items include item 3: "*I fully understand the patient's views regarding the important issues in making this decision";* item 11: *I am satisfied with the decision that was made*; item 12: *I am satisfied with the process used to make the decision was as good as it could be*.

Response options are rated on a scale of 1 to 5 where 5 shows "strongly agree" and 1 shows "strongly disagree". Items 6, 5, 4, 2, and 1 are negatively scored hence, they are reversed before calculating the total score of items 1-12. The maximum possible score range lies between 12 to 60. The scale has a sufficient alpha value ($\alpha = 0.87$). Participants were instructed via the demographic information sheet to recall a recent critical case they had managed. They were then asked to respond to the scale items based on their perception of that specific medical case.

Two additional measures were selected that were expected to be associated with decision-making; a team-decision-making questionnaire (Batorowicz & Shepherd, 2008) that measures the overall team support of healthcare providers in decision-making. This scale consists of 12 items and utilizes a seven-point Likert scale for rating. The response option for score 1 is "never" while the response option for score 7 is "to a large extent". The internal consistency of the measure is >.90 with all positively worded items. The scale is expected to correlate positively with PDPAI showing evidence of convergent validity. An additional sub-scale of professional quality of life i.e., compassion fatigue (CF) developed by Stamm and adapted by Galiana et al. (2020) was chosen to establish the discriminant validity of PDPAI. Compassion fatigue, also known as secondary or vicarious trauma, can be risk for healthcare providers who get exposed to patients' trauma, emergencies, and critical cases daily (Stamm, 2010). Physicians experience compassion fatigue because they deal with patients'trauma, emergencies, and serious cases daily. Chronic stress can result from the emotional toll of witnessing suffering coupled with the duties of the work. Compassion fatigue among healthcare personnel can be caused by a variety of factors, including long work hours, the difficult nature of medical decisions, and the pressure to strike a balance between empathy and professional detachment (Hui et al., 2023). It's a complex issue that goes beyond providing for patients'urgent needs to include the general emotional and mental health of individuals working in the medical field. This subscale employs a five-point Likert scale, with the option "Never" assigned a score of 1 and "Very Often" assigned a score of 5. The scale has a Cronbach alpha value of 0.82. It is expected that decision conflict will correlate positively with compassion fatigue.

Procedure

The study's aims were clearly explained during the recruitment procedure, and each participant gave their written informed consent before being included. Throughout the data collection procedure, confidentiality and privacy were prioritized. The necessary institutional review boards were consulted for ethical approval, which underlines the dedication to respecting ethical norms in research. They were given thorough information about the research including possible risks and benefits, and were given the assurance that participating was optional. To protect the participants'rights and wellbeing, all aspects of the process were conducted by ethical standards.

Participants were informed about the study's purpose, procedures, risks, and benefits. Written informed consent was obtained from all participants, ensuring their understanding and voluntary participation. Anonymity was ensured by using unique codes to protect participants'identities. Data was only accessible to the research team and was stored on password-protected computer. Additionally, to prevent individual identity, data was presented in aggregate form. Participants were told that data would be retained for a period of five years and then will be disposed of permanently deleting electronic files.

Data Analysis

SPSS version 26 was used to calculate the values of descriptive statistics. Group differences were assessed through independent sample T-test among the two physician groups. Multicollinearity diagnostics were assessed by calculating the Variance Inflation Factor (VIF) that showed values less than 5 indicating that multicollinearity was not an issue with the study variables. Moreover, we conducted the confirmatory factor analysis (CFA) through AMOS version 22 statistical software for three models (i) single-factor PDPAI (ii) two-factor PDPAI and (iii) bi- factor PDPAI to assess the factor structure of decision conflict. CFA is used to validate any existing measure because it helps in verification of the measure's factor structure with accuracy. Additionally, CFA is supportive in analysis of any underlying relationship between observed and latent factors. Fit indices were obtained with no residual dependency (Padgett & Morgan, 2021).

Factorial invariance of the PDPAI was examined across two groups of physicians' senior physicians (consultants/experts) (n = 167) and residents' trainees'/early career physicians (n = 180). Once the optimal factor solution was identified for both groups, the least restrictive configural invariance model was employed. By carefully investigating the distribution of fixed and free model parameters, this model evaluated the equivalence of the overall factor structure between the two groups without imposing any equality restrictions (Widaman & Olivera-Aguilar, 2023). To guarantee equality between groups, the metric invariance model subsequently set restrictions on each item's factor loading. The goal of this research was to ascertain whether the correlations between variables and factors were the same in both the early career residents in training and experienced physician groups. The aim was to investigate if every PDPAI item is consistently loaded into the same factor in both groups. Metric invariance, also known as weak measurement invariance, signifies consistent measurement units on the scale, indicating a shared understanding of the items among individuals in both groups. By limiting item intercepts, the scalar invariance model investigates if items share identical intercepts (item means) across both groups. Scalar invariance, commonly known as robust measurement invariance indicates that item scores in both groups adhere to a consistent measurement metric and identical scalar, enabling comparisons of factor means between the groups. The absence of scalar invariance implies the potential presence of systematic bias in response patterns between the two groups (Leroux et al., 2023). The stringent level of constraint is represented by the factor variance invariance model, commonly referred to as structural invariance.

To determine if the relationships between latent components are consistent across two groups of physicians, it sets extra limits on factor variances and covariances (Sass & Schmitt, 2013). This degree of invariance examines if the PDPAI scores for the two groups fall within the same range and whether the connection between the components is constant across physicians'groups.

Goodness of fit for the bi-factor model structure through structural invariance, factor variance invariance, metric invariance, and configural invariance were examined using these fit indices (i) Root Mean Square error of Approximation (RMSEA) (i) Comparative Fit Index (CFI) (iii) Standardized Root Mean Square Residual (SRMR). The subsequent cut-off values were employed as indicators of a well-fitting model; (1) RMSEA < .08 (2) CFI > 0.95 (3) SRMR < .08 (Hu & Bentler, 1999).

The present study also applied item response theory through Rasch modelling (Wright & Stone, 1999) on Jamovi analysis software program. An analysis of the outfit mean square (OUTFIT MNSQ) and infit mean square (INFIT MNSQ) statistics was undertaken (Smith, 2001). The analyses examine and assess the degree to which the observed data and the Rasch model's expected values correspond (Smith, 2000). This makes it easier to assess how much each component defines a shared construct. As suggested by Wright (1994), we deemed items to be "fit" if their MNSQ is between 0.6 and 1.4. Items with fit statistics between 0.5 and 1.5, however, may still be regarded as useful for measurement. A low score denotes simpler observations or less diversity in response patterns, while a high score denotes difficult observations or a wide range of responses to an item.

Additionally, discriminant validity was determined through Average Variance Extracted (AVE) and Maximum Shared Variance (MSV) criteria (Voorhees et al., 2016). The average variance of a factor should be higher than the average variance that it shares with all other factors (Farrell et al., 2009). To put this another way, the maximum shared variance (MSV) should be lower than the average variance (AVE) extracted. Also, the convergent validity was established through Pearson Product Moment correlation. PDPAI was correlated with team decision- making scale to assess convergent validity whereas the discriminant validity was assessed by correlating PDPAI with compassion fatigue. Since it is expected that PDPAI scale would positively correlate with team decision-making scale however lack of strong correlation is expected between PDPAI and compassion fatigue. The PDPAI's observed distinctiveness in the correlation pattern with its target construct, along with a modest link with compassion fatigue, are consistent with the findings of Fernández-Miranda et al. (2023) reinforcing the discriminant validity of the scale.

Finally, the Cronbach alpha was calculated for the overall sample and subscales as well as across two physician groups to determine the internal consistency reliability using SPSS software. In addition, McDonald Omega estimation (Ω) was calculated using JASP software freely available from the University of Amsterdam. Also, split-half reliability was calculated using the Spearman brown formula.

Results

Descriptives

Descriptive statistics are calculated for mean (*M*), standard deviation (*SD*), skewness (skew), and kurtosis (ku) for the scale. The values for conflict domain are (*M* = 16.06, *SD* = 4.10, skew = 0.13; ku = -0.11) while the values for satisfaction domain are (*M* = 25.75, *SD* = 4.03; skew= 0.06; ku = -0.07). Skewness and kurtosis lie within the acceptable range of ±1 showing the data is normally distributed.

Tab. 1 - *Group Differences on PDPAI scale and subscales*

Variables	Residents Trainee physicians		Consultant Physicians		t	p	95% CI		Cohen's d
	M	SD	M	SD			LL	UL	
PDPAI	41.09	5.99	43.22	6.41	-3.14	.00	-3.45	-8.08	0.34
D_Conflict	14.98	4.19	15.86	4.97	-1.73	.08	-1.35	.11	0.19
D_Satisfaction	26.10	4.17	27.35	4.09	-2.78	.00	-2.14	-3.68	0.30

Note. PDPAI = Provider Decision process Assessment Instrument, D_conflict = Decision conflict, D_Satisfaction = Decision satisfaction

Consultant physicians scored high on overall PDPAI scale as well as on decision satisfaction domain. Furthermore, non-significant differences are observed on decision conflict domain among the two groups (Table 1) Cohen's d is also too small and has almost negligible effect size. However, a modest effect size is indicated by a Cohen's d of 0.34 and 0.30 in case of group differences across overall scale and satisfaction domain respectively. This indicates that although there is a discernible difference between the two groups, it is not very significant. While the effect sizes are statistically significant, it is important to assess their practical significance. In practical terms, depending on the situation, even slight variations can have significant implications. For instance, small adjustments made to decision-making procedures may eventually result in better patient outcomes in a healthcare context. If there are not many differences between the groups, there may not be much of an impact on practice or policy. Stakeholders ought to contemplate whether the discernible disparities result in feasible modifications or enhancements that warrant the necessary exertion of time or resources. Future studies could explore these differences with larger sample sizes or different methodologies to verify whether the modest effect sizes persist or if more pronounced differences emerge. Researchers must also consider any additional variables or factors such as social and environmental variables that have an impact on effect sizes. Investigating these additional variables can offer a more thorough comprehension of the noted variations and their consequences.

Single Group CFA models

The measurement indices revealed a poor fit for the unidimensional decision conflict domain for either group of physicians. The values of fit indices were far from the acceptable range (AGFI = .35; GFI = .46; RMSEA = .27; IFI = .47; TLI = .35; CFI = .47). The two-factor structure of PDPAI comprising of "satisfaction" and "conflict" factors, demonstrated satisfactory fit in both early career physicians (CFI = .91, RMSEA = .08, SRMR = .05) and experienced physicians (CFI = .92, RMSEA = .04, SRMR = .06) groups. While this model improved on the unidimensional model, it still did not fully encapsulate the complexity of the PDPAI. Finally, the bifactor model, encompassing a singular general factor and two specific domain factors (conflict and satisfaction), demonstrated a good fit for both groups (CFI = .95, RMSEA = .05, SRMR = .03). As the bi-factor model was the only one exhibiting a satisfactory fit for both physician groups, examinations of structural and measurement invariance were solely performed on the bi-factor model. Significant standardized factor loadings were observed in the overall sample for the bi-factor model, ranging from 0.51 to 0.60 for the "conflict" factor and 0.44 to 0.70 for the "satisfaction" factor (Table 2).

Tab. 2 - *Factor loadings from Trainee Residents and Specialists baseline models of the PDPA*I

PDPAI-items	Postgraduate-Trainee Residents (*n* = 180)		Consultants (Experts) (*n* = 167)	
	Unstandardized	Standardized	Unstandardized	Standardized
Conflict				
1.	1.00	.58	1.00	.60
2.	1.08*	.54	1.02*	.55
4.	1.06*	.51	1.03*	.56
5.	0.76*	.51	0.78*	.60
6.	0.74*	.55	0.75*	.49
Satisfaction				
3.	1.00	.44	1.03	.50
7.	1.43*	.56	1.11*	.55
8.	1.20*	.58	0.89*	.60
9.	1.69*	.53	1.55*	.60
10.	0.99	.58	1.23	.66
11.	1.11	.67	1.00	.64
12.	1.00	.63	0.98	.70

Note: To establish the measurement scale for the underlying latent variable, the factorial loading of the first item was constrained to a value of 1,
*p < .05

Multigroup CFA models

We examined the fit indices of the factor variance invariance, scalar, metric, and configural, models across decision making of experienced physicians (consultants) and early career physicians (residents' trainees) for the bi-factor structure of the PDPAI which is presented in Table 3. For invariance testing, a change of less than or equal to 0.01 indicates that the invariance holds in ΔCFI. A change of less than or equal to 0.015 is considered acceptable for ΔRMSEA. For metric invariance, a change of less than or equal to 0.03 is acceptable for ΔSRMR; for scalar invariance, a change of less than or equal to 0.01 is acceptable. Additionally, for model fit (CFI ≥ 0.95, RMSEA ≤ 0.06, SRMR ≤ 0.08, TLI ≥ 0.95) indicates that the same factor structure is valid across groups. For metric invariance, overall model fit is reassessed and in comparison, to configural model using ΔCFI (≤ 0.01) and ΔRMSEA (≤ 0.015) is considered acceptable. Minimal changes in fit indices suggest that factor loadings are equivalent across groups. For scalar invariance, overall model fit is again reassessed compared to metric model using ΔCFI (≤ 0.01) and ΔRMSEA (≤ 0.015). For factor variance invariance, overall model fit is reassessed; compared to scalar model using ΔCFI (≤ 0.01) and ΔRMSEA (≤ 0.015). Minimal changes in fit indices suggest that factor variances are equivalent across groups.

Configural invariance

Without imposing any equality constraints on the two factor structure of model, configural invariance was investigated for the experienced and early career physicians groups. For both groups, the values of CFI, IFI, and SRMR were in acceptable range showing that the baseline two-factor model adequately fits the data. Statistically significant unstandardized factor loadings are obtained for both groups that are also in the same direction (λs = 0.74–1.69, *ps* < 0.001). This provides additional support of the bi-factor structural model across configural invariance.

Metric invariance

To assess metric invariance, all factor loadings were constrained to be equivalent across the two groups. The findings indicate that factor loadings remained consistent across the two groups, supported by a well-fitted model for the data (IFI > 0.90, CFI > 0.90, SRMR and RMSEA < 0.08) hence, metric invariance was achieved. The next step was to compare the less restrictive configural invariance model to the metric invariance. Consequently, results revealed descriptive (all Δ values <0.01) or non-statistical ($p > .05$) differences in model fit. Hence, the weak measurement invariance was satisfied and metric invariance model was considered a more fitting match for the data.

Scalar invariance

We examined the scalar invariance by constraining item intercepts and factor loadings to equivalence across the two groups. Both the item intercepts and factor loadings were invariant across the two groups as model fit was obtained for the scalar invariance (IFI > 0.90, CFI > 0.90, SRMR and RMSEA < 0.08). While making a comparison of the constrained model to the less restrictive metric invariance model, non-statistical descriptive differences (all Δ values < 0.01) were observed in model fit. Hence, there was a significant measurement invariance.

Factor variance invariance

We investigated the factor variance invariance by applying equivalence constraints across the two groups on item intercepts, factor loadings, and factor variance invariance. Acceptable model fit was obtained as revealed through the values of IFI, CFI and RMSEA. While comparing it to the less restrictive model of scalar invariance, a decline in model fit was observed (ΔCFI = -.016, ΔRMSEA = -.002, ΔSRMR = -.008). This indicates a dearth of factor variance invariance. Hence, the model with scalar invariance was deemed to be more optimal fit for the data.

Tab. 3 - *Goodness-of-fit indices for models testing scalar, configural, metric, and factor variance invariance of the PDPAI*

Model	Df	P	CFI	RMSEA	SRMR	Δdf	Δ p	Δ CFI	Δ RMSEA	Δ SRMR
Configural	104	< .001	.930	.073	.071	---	---	---	---	---
Metric	105	< .001	.920	.062	.075	1	.979	-.001	-.011	.004
Scalar	118	< .001	.930	.061	.078	13	.395	.01	-.001	.003
Factor Variance Invariance	120	< .001	.924	.059	.070	2	< .001	-.016	-.002	-.008

Rasch Analysis

The findings of the Rasch analysis suggested a generally favorable match (both infit and outfit) between the observed and model expected data (Table 4). The most challenging item found by participants was PDPAI item 3 ("I *was clear what treatment would be best for this patient*"), while item 7 was found to be the easiest one. Furthermore, person reliability statistics revealed 0.824 which is a good fit. In our Rasch analysis, items 1, 2, 3, 4, 5, and 6 were identified as difficult. Items with poor fit have the potential to skew measurements and diminish the instrument's validity. Significantly misfitting items may not be appropriate to all respondents or may measure something different from the intended concept. However, A DIF analysis was performed to check whether these items displayed variations in difficulty or discrimination between subgroups. This analysis helped us ensure that the items are functioning fairly across different groups. No significant DIF was detected in these difficult items (1, 2, 4, 5, 6), indicating that they may not be equally challenging or discriminative for the two subgroups. If the items showed significant DIF, this could introduce bias, making it difficult to compare scores across the two groups. However, there are several reasons for misfit of the items. (i) Items could measure something different or less relevant than expected if they don't align well with the intended concept. The item content may not be clear, or the language may be confusing, causing this mismatch. (ii) Secondly, if respondents perceive the items differently, poor fit may result in inconsistent results. These biases may result from response styles that differ, cultural variations, or differing levels of familiarity with the item content. (iii) Thirdly, Misfit can occur when items with extremely high or low difficulty levels do not align well with the overall distribution of respondent skills. For example, an item may not fit well into the Rasch model if it is too simple or difficult in comparison to the other items. (iv) Lastly, the fit of some items may be impacted by their redundancy or excessive complexity. Items that are very complex or comparable to one another may not offer any new or distinctive insights into the construct being measured.

Furthermore, the Wright map is a very useful sort of statistic which has been outputted in Figure 1. On the right-hand side are the items which have been sorted based on their difficulty. For example, item no. 9, 10, 7, 8, 12, and 11 cluster together and are the easiest ones which fall at the bottom then there is a little bit of gap between these items and the rest of the items. On the left-hand side is the distribution of people which has been demonstrated using bar graphs. Hence, there is a good distribution in items as well as people. Since there are people from all levels of ability and the ability ranges from around -2 logits to the estimated ability of larger than 4. Here, it is also essential to mention that as per the person reliability statistic which is 0.824; indicates that we have been able to estimate the ability levels of our test takers with 82.4 percent of precision.

Tab. 4 - *MNSQ Infit and Outfit values indicating the conformity of the observed data to the model's expected data*

Items no.	MNSQ INFIT	MNSQ OUTFIT
Dec_PA1_1	0.995	0.980
Dec_PA2_1	1.087	1.099
Dec_PA3_1	1.327	1.278
Dec_PA4_1	1.118	1.103
Dec_PA5_1	1.099	1.106
Dec_PA6_1	1.129	1.133
Dec_PA7_1	0.902	0.869
Dec_PA8_1	0.891	0.872

Dec_PA9_1	0.941	0.925
Dec_PA10_1	0.892	0.922
Dec_PA11_1	0.937	0.942
Dec_PA12_1	1.121	1.100

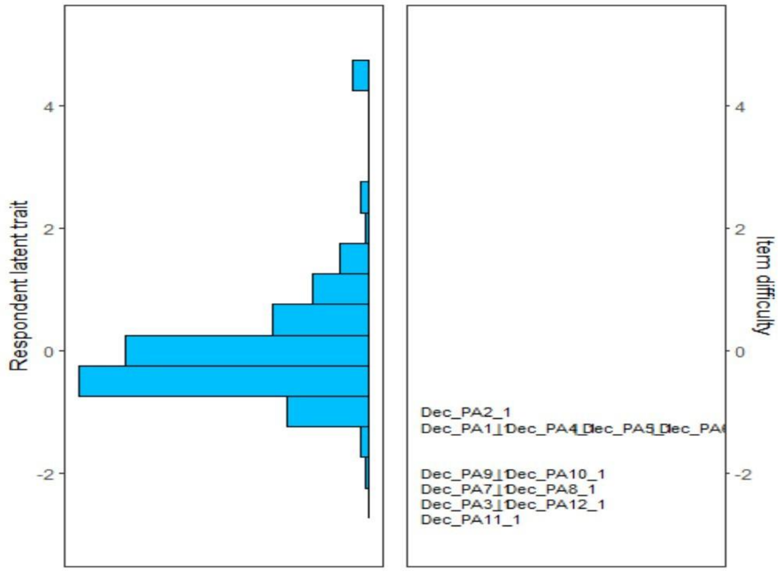

Fig. 1 - *A Graphical Representation of Item Response Theory through Wright Map*

Internal Consistency and McDonald's Omega Reliability

Internal consistency for the PDPAI scale was adequate for the total sample (α = .79), early career physicians' group (α = .87), and experienced physicians' group (α = .78). The factor 1 consisting of conflict items revealed alpha (α = 0.86) while Factor 2 consisting of satisfaction items revealed alpha (α = 0.81) for early career physicians (residents) while Factor 1 revealed alpha (α = 0.81) and factor 2 revealed alpha (α = 0.81) for experienced (consultants) physicians**.** The overall McDonald's omega reliability was found to be (**Ω** = 0.76), however for satisfaction items it was (**Ω** = 0.85) and for conflict items, the omega coefficient was (**Ω** = 0.81). Additionally, using the Spearman-Brown formula, which adjusts for the fact that we only used half of the scale at a time, we estimated the reliability of the full scale. The split-half reliability was found to be 0.77, which is satisfactory. This reliability indicates that the two halves of the scale are very consistent with each other.

Convergent and Discriminant validity

The value of AVE is found above the cut-off value of 0.50 (Henseler et al., 2016). The factor loadings of the items for each respective construct helped in determining the convergent validity. The percentage of AVE for both the factors; satisfaction (AVE = 0.511, MSV = 0.373) and decision conflict (AVE = 0.542, MSV = 0.440) was in acceptable range. The construct's individual correlation with each of the other factors was smaller than the square root of average variance extracted AVE values for each sub-scale. Considering this, it is suggested that the AVE of each individual factor was more than the MSV.

Moreover, we also established criterion-related validity by Pearson correlation. Decision satisfaction was correlated with team support in decision-making (r = .484**, p < .00) showing a positive correlation. The effect size is moderate. In practical terms, this suggests that change in decision satisfaction moderately predict change in team support. Additionally, decision conflict correlated positively with compassion fatigue (r = .382**, p < .00) with moderate effect size. Hence, evidence of convergent validity is supported. Additionally, decision satisfaction correlated negatively with compassion fatigue (r = -.099*, p <.05) showing evidence for discriminant validity. However, the strength of the effect size is almost negligible in size.

Discussion

A significant lacuna in the psychometric assessment of the decision assessment scale was identified by the literature, which paid scant attention to important factors including measurement invariance, convergent validity, and discriminant validity. To bridge this gap, the current study carried out a thorough measurement invariance test as well as a rigorous psychometric assessment using both single and multigroup Confirmatory Factor Analysis (CFA). Additionally, the investigation of convergent and discriminant validity deepens our comprehension of the robustness of the PDPAI and provides a nuanced contribution to the existing body of literature.

The PDPAI demonstrated a favorable fit within a sample comprising resident trainees and (experienced) consultant physicians, establishing itself as a reliable and valid measure for this particular demographic. Acceptance of a bi-factor model, alongside well-fitted factor variance invariance, scalar, metric, and configural equivalent models, underscores the appropriateness of the instrument for the dataset. The bifactor model affirms the suitability of employing either the subdomains or a global measurement model of the PDPAI.

Factorial invariance analysis on the two groups of resident trainees (early career) and consultants (experienced) physicians indicated that metric and scalar invariance showed a better fit than that of the configural or factor variance invariance models, suggesting that PDPAI can be applied to both experienced physicians and early career trainees. The factor variance invariance of the PDPAI was not achieved, and it was possible to distinguish the correlational associations between the latent variables of the two groups. The instrument remains valid for comparing the average levels of the latent construct between groups due to achieved scalar invariance. However, we are unable to compare the degree of variability or dispersion of the construct between groups due to the inability to demonstrate factor variance invariance. Comparing the latent construct's distribution or degree of variability between groups becomes difficult in the absence of factor variance invariance. Since the instrument does not measure variability consistently across groups, differences in construct variability cannot be evaluated with confidence. Future studies may consider looking at possible causes of group-to-group variability in the construct. This can entail determining if particular items contribute disproportionately to the variance or whether response pattern variations between groups have an impact on the dispersion of the concept. Furthermore, future researchers may look into different measurement models or methods that can provide a better alignment of factor variation between groups. This can entail improving the instrument's fairness and usefulness for a variety of respondent groups.

Although the structural invariance was not present, we reemphasize that the scores within the PDPAI were appropriate for different levels of physicians (early career residents and more experienced consultants). Measurement invariance (both metric and scalar) is thought to be essential for cross-group comparisons. Cheung (2004) describes invariance as 'measuring the same construct' as similar scores between the two groups are less likely to be a result of sample bias, and more likely to represent true group differences.

Rasch analysis confirmed a good fit between empirical and expected data and are in line with convincing evidence from current research (e.g., Paceco-Colón et al., 2019). This validates the importance of accuracy in measuring instruments to capture the variability that exists in decision-making processes. The problems mentioned, such as, respondents' inability to define what would be best for patients (PDPAI question 3), complexity of their decision process (PDPAI question 2) and easiness to find aspects to influence the decision (PDPAI item 7), were previously found in the relevant literature. Zhou and Xu (2023) found that medical decision-making is 'an extremely complex process at its core'. Moral distress in relation to clinical ethics and end-of-life decision-making is a significant problem in healthcare practice and a very distressful experience for those involved. The importance of cognitive fluency for decision-making was found by Schwarz and colleagues in 2021. Taking into consideration the aforementioned evidence, and its robustness, it is possible to assume that each of the answered items of the PDPAI revealed even subtle and very small shifts in decision-making processes among physicians. This finding demonstrates one of the PDPAI's main implications to enhance the practical and field research as well as to strengthen its psychometric properties. Another implication refers to potential improvements in interventions created to tackle and alleviate decisional conflict among cancer patients and their families.

Furthermore, the internal consistency of the PDPAI scale revealed substantial alpha coefficients in both groups of the physicians as well as in the total sample. Strong McDonald's omega reliability and consistency in the factor structures for the satisfaction and conflict items reinforce the internal coherence of the scale. These findings demonstrate how well the PDPAI captures the nuanced nature of healthcare professionals'decision-making experiences.

The study found each individual factor's AVE was higher than the MSV, supporting discriminant validity and indicating that each construct is assessing a distinctive and particular component of the phenomenon under study. Thus, the claim that the measurement tool effectively distinguishes between various conceptions is supported by the fact that items inside a construct are more likely to agree than things from other constructs. This study strengthens the measurement model's validity by showing that the constructs are sufficiently distinct and do not overly overlap, adding to the assessment tool's overall robustness. Moreover, criterion-related validity is confirmed by Pearson correlations. The correlation between decision satisfaction and team support highlights the link between successful decision-making and collaborative settings (Wang et al., 2023). The positive relationship between decision conflict and compassion fatigue also emphasizes how decisional difficulties affect caregiver wellbeing (Cocker & Joss, 2016). Positive decision experiences may help to lessen the effects of caregiver fatigue, according to the negative association between decision satisfaction and compassion fatigue. These results confirm the validity of the assessment technique by highlighting the complex interactions between decision-making, team support, and compassion fatigue.

Limitations of the Study

We used convenience sampling which limits the generalizability of the study findings. This is especially true for physicians who practice in different areas, in different types of healthcare facilities, or from different cultural backgrounds. It is critical to consider our results in light of this sampling strategy. We contrasted our results with those from other studies (e.g., Wu et al., 2022) that used similar sampling strategies to lessen this constraint. This comparative study showed that, despite the sampling limitations, our results are consistent with more general trends observed in the field. However, more robust sampling methods, such as random or stratified sampling are recommended for future researchers to enhance the generalizability of the findings. Moreover, follow-up studies using different methodologies (for instance, focus groups or qualitative interviews) could help expand and validate our initial findings.

The study did not explicitly account for cultural factors in the analysis, which is a limitation that needs to be addressed in future research. Cross-cultural studies should be conducted to assess the validity and reliability of the PDPAI. These studies should include diverse samples to capture the cultural variability in decisional conflict. Modification of items as per the cultural relevance is essential which may involve adding culturally specific examples, rephrasing of items, or adjusting response options. Moreover, cultural variables could be included in the analysis to examine how cultural differences impact decisional conflict. This could involve comparing results across different cultural groups and identifying culturally specific patterns.

Conclusion

The PDPAI emerges as a reliable and valid measure for usage among healthcare physicians. The overall measure of decisional conflict, the sub-domains of the scale, and individual items all demonstrated a satisfactory congruence with the collected data. Evaluating decisional conflict stands as a crucial endeavor to gauge the challenges encountered by health physicians in the decision-making process. The PDPAI can be used in future research projects aiming at improving decision-making processes to examine if such improvements help in reduction of decision conflicts among physicians.

Practical Implications

Enhancing the decision-making process by reducing decisional conflict presents a significant problem for physicians. The practicality of totally eliminating uncertainty was criticized by Nelson and colleagues (2007) who said, “for several decisions, the objective of entirely eliminating or even reducing uncertainty is merely impractical" (p. 615). However, it is crucial to understand that within the framework of the decisional conflict, uncertainty just represents one aspect of decisional conflict (O'Conner, 1995). Healthcare professionals can address different subdomains of decisional conflict notwithstanding the inherent ambiguity in difficult medical decisions. They can do this by educating patients, offering assistance, and soliciting their values in an effort to reduce medical decisional conflict.

It is crucial to take into account the effects of decisional conflict in the context of healthcare decision-making procedures given the potential difficulties faced by healthcare professionals who are involved in decision-making. Evaluating decisional conflict as a key indicator is advisable given the complex dynamics and pressures involved in making important decisions for patients. An analogous consideration for what might be referred to as a healthcare professional's decisional dilemma – a situation in which the complexity and importance of healthcare decisions place a heavy load on the decision-makers – exists, building on the idea Decisional conflict is a crucial metric for assessing interventions supporting healthcare professionals in navigating complex healthcare decision-making. Drawing from Netzer and Sullivan (2014), measuring decisional conflict illuminates the challenges faced by professionals, providing a key gauge for intervention effectiveness.

References

Barry, M. J., Cherkin, D. C., YuChiao, C., Fowler, F. J., & Skates, S. (1997). A randomized trial of a multimedia shared decision-making program for men facing a treatment decision for benign prostatic hyperplasia. *Disease Management and Clinical Outcomes*, *1*(1), 5-14.

Batorowicz, B., & Shepherd, T. A. (2008). Measuring the quality of transdisciplinary teams. *Journal of Interprofessional Care*, *22*(6), 612-620. DOI: 10.1080/13561820802303664.

Bauer, D. J. (2017). A more general model for testing measurement invariance and differential item functioning. *Psychological Methods*, *22*(3), 507-526. DOI: 10.1037/met000007.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233-255.

Cocker, F., & Joss, N. (2016). Compassion fatigue among healthcare, emergency and community service workers: A systematic review. *International Journal of Environmental Research and Public Health*, *13*(6), 618.

Dhami, M. K., & Mandel, D. R. (2022). Communicating uncertainty using words and numbers. *Trends in Cognitive Sciences*, *25*(6) 514-526 DOI: 10.1016/j.tics.2022.03.002.
Dolan, J. G. (1999). A method for evaluating health care providers'decision making: the Provider Decision Process Assessment Instrument. *Medical Decision Making*, *19*(1), 38-41.
Epstein, S. (2011). Intuition from the perspective of cognitive-experiential self-theory. In Plessner H., Betsch C., & Betsch T. (Eds.), *Intuition: In judgment and decision making* (pp. 23-37). Hoboken: Taylor & Francis.
Falzer, P. R. (2018). Naturalistic decision making and the practice of health care. *Journal of Cognitive Engineering and Decision Making*, *12*(3), 178-193. DOI: 10.1177/1555343418773915.
Farrell, A. M., & Rudd, J. M. (2009). *Factor analysis and discriminant validity: A brief review of some practical issues*. Anzmac.
Fernández-Miranda, G., Urriago-Rayo, J., Akle, V., Noguera, E., Mejía, N., Amaya, S., & Jimenez-Leal, W. (2023). Compassion and decision fatigue among healthcare workers during COVID-19 pandemic in a Colombian sample. *Plos One*, *18*(3), e0282949.
Galiana, L., Oliver, A., Arena, F., De Simone, G., Tomás, J. M., Vidal-Blanco, G., ... & Sansó, N. (2020). Development and validation of the Short Professional Quality of Life Scale based on versions IV and V of the Professional Quality of Life Scale. *Health and Quality of Life Outcomes*, *18*(1-12). DOI: 10.1186/s12955-020-01618-3.
García, A., & Garasic, M. (2021). Ethical issues concerning informed consent in translational/clinical research and vaccination. In *Cross-Cultural and Religious Critiques of Informed Consent*,(pp. 11-17). Routledge.
Hair, J. F., Anderson, R. E., Babin, B. J., & Black, W. C. (2010). *Multivariate data analysis: A global perspective*, (7). Pearson Upper Saddle River.
Han, P. K., Babrow, A., Hillen, M. A., Gulbrandsen, P., Smets, E. M., & Ofstad, E. H. (2019). Uncertainty in health care: Towards a more systematic program of research. *Patient Education and Counseling*, *102*(10), 1756- 1766.
Hong, S., Malik, M. L., & Lee, M. K. (2003). Testing configural, metric, scalar, and latent mean invariance across genders in sociotropy and autonomy using a non-Western sample. *Educational and Psychological Measurement*, *63*(4), 636-654. DOI: 10.1177/0013164403251332.
Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: a Multidisciplinary Journal*, *6*(1), 1-55. DOI: 10.1080/10705519909540118.
Hui, L., Garnett, A., Oleynikov, C., & Boamah, S. A. (2023). Compassion fatigue in healthcare providers during the COVID-19 pandemic: a scoping review protocol. *BMJ open*, *13*(5), e069843.
Klein, G. (2015). A naturalistic decision making perspective on studying intuitive decision making. *Journal of Applied Research in Memory and Cognition*, *4*(3), 164-168. DOI: 10.1016/j.jarmac.2015.07.001.

Légaré, F., O'Connor, A. M., Graham, I. D., Wells, G. A., & Tremblay, S. (2006). Impact of the Ottawa Decision Support Framework on the agreement and the difference between patients'and physicians'decisional conflict. *Medical Decision Making*, *26*(4), 373-390.

Leroux, E. J., Rizeq, J., & Skilling, T. A. (2023). Measurement invariance of the Youth Self-Report across youth who have committed sexual and nonsexual offenses. *Psychological Assessment*, *35*(10), 821-829.

Liao, Y. L., Wang, T. J., Su, C. W., Liang, S. Y., Liu, C. Y., & Fan, J. Y. (2023). Efficacy of a decision support intervention on decisional conflict related to hepatocellular cancer treatment: a randomized controlled trial. *Clinical Nursing Research*, *32*(1), 233-243.

Liu, Y., Wang, X., Wang, Z., Zhang, Y., & Jin, J. (2023). Ethical conflict in nursing: A concept analysis. *Journal of Clinical Nursing*, *32*(15-16), 4408-4418. DOI: 10.1111/jocn.16563

Moure, M., Jacobsen, J. B., & Smith-Hall, C. (2023). Uncertainty and climate change adaptation: a systematic review of research approaches and people's decision-making. *Current Climate Change Reports*, *9*(1), 1-26. DOI: 10.1007/s40641-023-00189-x

Nelson, W. L., Han, P. K., Fagerlin, A., Stefanek, M., & Ubel, P. A. (2007). Rethinking the objectives of decision aids: a call for conceptual clarity. *Medical Decision Making*, *27*(5), 609-618. DOI: 10.1177/0272989X07306780.

Netzer, G., & Sullivan, D. R. (2014). Recognizing, naming, and measuring a family intensive care unit syndrome. *Annals of the American Thoracic Society*, *11*(3), 435-441.

O'Connor, A. M. (2010). *User Manual-Decisional Conflict Scale. 1993.* See [https://decisionaid. ohri. ca/eval. html# DecisionalConflictScale].

O'Donnell, D., O'Donoghue, G., Ní Shé, É., O'Shea, M., & Donnelly, S. (2023). Developing competence in interprofessional collaboration within integrated care teams for older people in the Republic of Ireland: A starter kit. *Journal of Interprofessional Care*, *37*(3), 480-490. DOI: 10.1080/13561820.2022.2075332.

O'Connor, A. M. (1995). Validation of a decisional conflict scale. *Medical Decision Making*, *15*(1), 25-30. DOI: 10.1177/0272989X9501500105.

O'Connor, A. M., Jacobsen, M. J., & Stacey, D. (2002). An evidence-based approach to managing women's decisional conflict. *Journal of Obstetric, Gynecologic, & Neonatal Nursing*, *31*(5), 570-581.

Pacheco-Colón, I., Hawes, S. W., Duperrouzel, J. C., Lopez-Quintero, C., & Gonzalez, R. (2019). Decision-making as a latent construct and its measurement invariance in a large sample of adolescent Cannabis users. *Journal of the International Neuropsychological Society*, *25*(7), 661-667. DOI: 10.1017/S1355617719000341.

Padgett, R. N., & Morgan, G. B. (2021). Multilevel CFA with ordered categorical data: A simulation study comparing fit indices across robust estimation methods. *Structural Equation Modeling: A Multidisciplinary Journal*, *28*(1), 51-68. DOI: 10.1080/10705511.2020.1759426.

Pecanac, K. E., Brown, R. L., Steingrub, J., Anderson, W., Matthay, M. A., & White, D. B. (2018). A psychometric study of the decisional conflict scale in surrogate decision makers. *Patient Education and Counseling*, *101*(11), 1957-1965. DOI: 10.1016/j.pec.2018.07.006.

Ruzsa, G., Szeverenyi, C., & Varga, K. (2020). Person-and job-specific factors of intuitive decision-making in clinical practice: results of a sample survey among Hungarian physicians and nurses. *Health Psychology and Behavioral Medicine*, *8*(1), 152-184. DOI: 10.1080/21642850.2020.1741372.

Sainfort, F., & Booske, B. C. (2000). Measuring post-decision satisfaction. *Medical Decision Making*, *20*(1), 51-61.

Sass, D. A., & Schmitt, T. A. (2013). Testing measurement and structural invariance: Implications for practice. In *Handbook of quantitative methods for educational research* (pp. 315-345). Brill.

Scholl, I., Koelewijn-van Loon, M., Sepucha, K., Elwyn, G., Légaré, F., Härter, M., & Dirmaier, J. (2011). Measurement of shared decision making-a review of instruments. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen*, *105*(4), 313-324. DOI: 10.1016/j.zefq.2011.04.012.

Schwarz, N., Jalbert, M., Noah, T., & Zhang, L. (2021). Metacognitive experiences as information: Processing fluency in consumer judgment and decision making. *Consumer Psychology Review*, *4*(1), 4-25.

Simon, D., Loh, A., & Härter, M. (2007). Measuring (shared) decision-making-a review of psychometric instruments. *Zeitschrift für ärztliche Fortbildung und Qualität im Gesundheitswesen-German Journal for Quality in Health Care*, *101*(4), 259-267. DOI: 10.1016/j.zgesun.2007.02.029.

Smith Jr, E. V. (2001). Evidence for the reliability of measures and validity of measure interpretation: a Rasch measurement perspective. *Journal of Applied Measurement*, *2*(3), 281-311.

Smith, R. M. (2000). Fit analysis in latent trait measurement models. *Journal of Applied Measurement*, *1*(2), 199-218.

Strauss, M. E., & Smith, G. T. (2009). Construct validity: Advances in theory and methodology. *Annual Review of Clinical Psychology*, *5*, 1-25.

Trees, A. R., Ohs, J. E., & Murray, M. C. (2017). Family communication about end-of-life decisions and the enactment of the decision-maker role. *Behavioral Sciences*, *7*(2), 36.

Voorhees, C. M., Brady, M. K., Calantone, R., & Ramirez, E. (2016). Discriminant validity testing in marketing: an analysis, causes for concern, and proposed remedies. *Journal of the Academy of Marketing Science*, *44*, 119-134. DOI: 10.1007/s11747-015-0455-4.

Wang, S., Li, L., Liu, C., Huang, L., Chuang, Y. C., & Jin, Y. (2023). Applying a multi-criteria decision-making approach to identify key satisfaction gaps in hospital nurses' work environment. *Heliyon*, *9*(3)e14721.

Wendler, D., & Rid, A. (2011). Systematic review: the effect on surrogates of making treatment decisions for others. *Annals of Internal Medicine*, *154*(5), 336-346. DOI: 10.7326/0003-4819-154-5-201103010-00008.

Widaman, K. F., & Olivera-Aguilar, M. (2023). Investigating measurement invariance using confirmatory factor analysis. *Handbook of structural equation modeling*, 367-384.
Wilson, M., Allen, D. D., & Li, J. C. (2006). Improving measurement in health education and health behavior research using item response modeling: comparison with the classical test theory approach. *Health Education Research*, *21*(suppl_1), i19-i32.
Wright, B. D. (1994). Reasonable mean-square fit values. *Rasch Meas Transac*, *8*, 370.
Wright, B. D., & Stone, M. H. (1999). *Measurement Essentials*. Wilmington, DE: Wide Range.
Wu, D., Yang, T., Herold, F., Hall, D. L., Mueller, N., Yeung, A., ... & Zou, L. (2022). Validation of the 4-Item and 10-Item uncertainty stress scale in a community-based sample of chinese adults. *Psychology Research and Behavior Management*, 2803-2813.
Younas, S. & Khanum, S. (2024a). Examining the Role of Stress and Team Support in Decision-Making Under Uncertainty and Time Pressure. *MDM Policy and Practice* (in press).
Younas, S. & Khanum, S. (2024b). Unraveling the Predictive Role of Work Rules on Compassion Satisfaction and Career Satisfaction among Professionals of Obstetrics and Gynaecology: The Mediating Effect of Team Support. *Journal of Professions and Organization*, *11*(3)1-15. DOI: 10.1093/jpo/joae011.
Younas, S., Khanum, S., & Qamar, A. H. (2023). Decision making among residents in training of obstetrics and gynecology: A qualitative exploration in Pakistani context. *PLoS ONE 18*(11), e0287592. DOI: 10.1371/journal.pone.0287592.
Zettel-Watson, L., Ditto, P. H., Danks, J. H., & Smucker, W. D. (2008). Actual and perceived gender differences in the accuracy of surrogate decisions about life-sustaining medical treatment among older spouses. *Death Studies*, *32*(3), 273-290.
Zettel-Watson, L., Ditto, P. H., Danks, J. H., & Smucker, W. D. (2008). Actual and perceived gender differences in the accuracy of surrogate decisions about life-sustaining medical treatment among older spouses. *Death Studies*, *32*(3), 273-290. DOI: 10.1080/07481180701881230.
Zhou, J., & Xu, X. (2023). The difficulty of medical decision-making: should patients be involved?. *Hepatobiliary Surgery and Nutrition*, *12*(3), 407-409.