CHAPTER-3

METHODOLOGY

3.0 Introduction

The introduction provided an in-depth summary of the research's context, significance, necessity, aims, boundaries, and restrictions. The second chapter focused on examining related literature and conducting research reviews. In the third chapter, we elaborated on the research methodology and procedure, which comprise four components: research method, sampling method, tool development, and research implementation.Research methodology is the methodical, scientific method that scientists employ to collect and examine data to find answers to research questions or validate hypotheses. It is a critical component of any research project, providing the context for assembling and understanding data.

Research methodology is crucial for ensuring that the study is carried out thoroughly and systematically, resulting in dependable and accurate outcomes. One of the key benefits of research methodology is that it provides a clear and structured approach to carrying out research. It helps researchers identify the appropriate research proposal, sample techniques, and information collection methods. This ensures that the research is conducted consistently and standardized and that the data collected is reliable and valid. This is essential for maintaining that the research results are accurate and can be used to make informed decisions.

Research methodology also helps researchers recognize biases and errors in their research. Using appropriate research methods and techniques, researchers can minimize the impact of these biases and errors and ensure that the research findings are as accurate and reliable as possible. Research methodology also helps to ensure that research is conducted ethically. Researchers must adhere to strict ethical guidelines when performing research, and research methodology provides a framework for ensuring these guidelines are observed. Securing informed agreement from all participants, ensuring the confidentiality of their information, and protecting their rights and well-being are critical aspects of ethical research conduct. In addition, research methodology helps to ensure that research is conducted efficiently and

effectively. Researchers can minimize the time and resources required by using appropriate research methods and techniques while producing reliable and valid results. This is essential for ensuring that research is conducted cost-effectively and that resources are utilized efficiently.

Overall, research methodology is a vital constituent of any research project. It provides a clear and structured approach to conducting research, helps to identify potential biases and errors, ensures that research is conducted ethically, and enables researchers to conduct research efficiently and effectively. Without research methodology, research would be performed ad hoc, producing unreliable and invalid results. Research is a systematic and scientific approach to gathering data, analyzing it, and drawing conclusions to answer research questions or test hypotheses.

3.1 Rationale of the study

The assessment of professional behavior is crucial for evaluating the success of teaching strategies in any institute. Without proper assessment, it's difficult to determine the effectiveness of educational approaches. The need for developing new assessment tools arises from the growing interest in teaching and evaluating professionalism. It's important to create culturally relevant assessment tools, especially in the context of dental education in India. Existing literature suggests that assessing professionalism in dentistry is time-consuming but essential. Therefore, there's a need to develop a professional assessment scale tailored to the Indian cultural and professional framework. This scale should integrate input from patients, dental experts, faculty, and students to confirm its reliability and validity. With this in mind, a study is proposed to create and validate a professional assessment scale for postgraduate dental students.

Dental education should place significant importance on the advancement of soft skills alongside the acquisition of theoretical knowledge. Professionalism stands out as a vital attribute that aspiring dental surgeons must diligently cultivate. A dentist who possesses robust professional skills is not only better prepared to navigate the complexities of clinical practice but also effectively manage their personal and professional life, fostering a holistic approach to their career.

Recognizing the need for professionalism is merely the first step; its active development is equally crucial. To elevate the standards of professionalism, it becomes essential to pinpoint the various factors that contribute to it. A key question in this exploration is how these elements can be assessed accurately. Research with in the realm of dental education is focused on exploring these complexities. It is evident that before we can implement strategies to enhance professionalism among dental students, we must first evaluate its current state.

The initial step in assessing professionalism involves identifying the potential components that constitute this multifaceted concept. Following this identification, it is important to develop a specialized assessment instrument grounded in these components, capable of measuring professionalism effectively. At present, there is no established scale in India tailored specifically for postgraduate dental students. Nonetheless, the investigator tried creating such an assessment tool aimed at measuring professionalism. Although some international scales exist in the literature, they often lack applicability to Indian students due to cultural and contextual differences.

The longstanding maxim that "assessment drives learning" holds particular significance in this context. The assessment framework employed by any educational institution mirrors the effectiveness of the teaching strategies implemented within it. Without a methodical assessment of professional behavior, educators cannot ascertain whether their pedagogical efforts have been fruitful. Professionalism encompasses a wide array of domains that are often challenging to evaluate. As Arnold famously stated, "Without reliable assessment tools, questions about the effectiveness of approaches to educating learners about professional behavior will not be effectively answered."

Therefore, a pressing need was felt to develop innovative assessment tools in response to the increasing interest in teaching and evaluating professionalism. Crafting assessment tools that resonate with the cultural nuances of dental education in India is particularly important. Existing literature indicates that while assessing professionalism in dentistry may be a time-intensive process, it is nonetheless indispensable. This scale should incorporate feedback from a diverse range of stakeholders, including patients, dental practitioners, faculty, and students, to ensure its successful application. In light of this need, a study is proposed to develop and validate a professional assessment scale tailored for postgraduate dental students, ultimately aiming to augment the quality of dental education and practice in India.

3.2 Statement of the problem

Standardization of Scale to measure professionalism in postgraduate dental students.

In India, the landscape of dental education is evolving, yet a significant gap persists concerning the assessment and development of professionalism in postgraduate dental students. Professionalism is a multifaceted construct that encompasses not only clinical competence but also adherence to ethical standards, communication skills, and an understanding of societal values. Despite the critical role professionalism plays in ensuring high-quality patient care and fostering trust within the healthcare system, there is currently no standardized framework that articulates the essential competencies expected of dental professionals in the Indian context.

Existing international frameworks for measuring professionalism often do not account for the unique societal and cultural nuances of India. They may overlook important dimensions such as the interplay between personal and professional identity, the ethical obligations towards diverse populations, and the expectations of various stakeholders including patients, educators, and the community at large. This lack of context-specific guidance leads to inconsistencies in how professionalism is understood and taught in dental programs, ultimately impacting the preparedness of graduates to meet the psychological needs of society.

Furthermore, the absence of a comprehensive assessment tool means that important components of professionalism may be neglected in contemporary dental education, leading to potential deficits in important skills such as empathetic communication, ethical decision-making, and teamwork. As a result, postgraduate dental students may graduate without a fully developed professional identity or the competence required to navigate complex ethical dilemmas in their practice.

In light of these challenges, this study strives to address the urgent need for a tailored framework to measure professionalism among postgraduate dental students in India. By identifying the foundational values and competencies essential for professionalism within the Indian context, this research aims to not only facilitate better assessment but also to inform the development of targeted educational strategies that foster these ideals within dental programs. Ultimately, this effort will contribute to enhancing the quality of dental

education and practice in India, ensuring that graduating professionals are well-equipped to serve the needs of their communities.

3.3 Objectives of the Study

1. To analyze and identify the constructs and domains of professionalism for Indian postgraduate dental students.
2. To prepare and standardize the scale of professionalism for Indian postgraduate dental students.
3. To implement the scale of professionalism in the context of Indian postgraduate dental students to develop norms.
4. To study the relationship of the professionalism scale with various demographic factors.

3.4 In this study, the concept of professionalism was operationalized as

Professionalism is described as the integration of various essential skills of an individual to act ethically in both their personal and professional lives, while also advocating for and contributing to societal benefits

3.5 Hypotheses

Ho1: There will be no significant difference in the mean professionalism score of male and female students

Ho2: There will be no significant difference in the mean professionalism score of married and unmarried students

Ho3: There will be no significant difference in the mean professionalism scores of students of parents with health care or no healthcare backgrounds.

Ho4: There will be no significant difference in the mean professionalism score between students who opted by interest or by compulsion

Ho5: There will be no significant difference in the mean professionalism score between students having siblings or no siblings

Ho6: There will be no significant difference in the mean professionalism scores of students of urban and rural environments.

Ho7: There will be no significant differences in the mean professionalism scores of students according to I, II, and II years of postgraduation

Ho8: There will be no significant difference in the mean professionalism scores of students according to various Family incomes

Ho9: There will be no significant difference in the mean professionalism scores of students according to whether they attended any formal training /courses on professionalism and ethics or not

3.6 Delimitation of the Study

The research was delimited to postgraduate dental students studying in the dental colleges of India.

3.7 Data required for the research

The following data was required:

1. Professionalism scale score.
2. Family income status
3. Gender.
4. Marital status
5. Post graduation year
6. Opting the profession by interest or compulsion
7. Sibling status
8. Urban and rural background
9. Healthcare or non-healthcare background.

3.8 Source of Data: Dental Colleges in India offering postgraduate courses were included as sources of data. All three years of postgraduate students with nine branches of specialization willing to participate in the study comprised for the source of data.

3.9 Population:

The population consisted of all postgraduate dental students in India who were enrolled in various dental colleges offering postgraduate courses. There are approximately 200 colleges in India that provide doctoral programs, but the number of available seats is subject to change every year based on annual inspections and the availability of faculty. Additionally, not all seats are filled in either private or government dental colleges.

Sample:

In this study, we aimed to develop and standardize a tool to measure the professionalism of postgraduate students. This was the first phase of the research, which required two separate sample groups for standardization and subsequent implementation of the tool.

Phase-1 sample

Pilot study with 100 postgraduate students

Data was collected between December 2020 and August 2023. Due to the significant restrictions during the COVID-19 pandemic, the standardization phase of the scale began under challenging conditions. The researcher was only able to conduct a pilot study through direct physical interaction at two dental colleges with postgraduate programs in Vadodara.

In the first phase of constructing the tool, 100 postgraduate students participated in a pilot study. For the pilot study, all post-graduates from two dental colleges in Vadodara were selected**.**

College Name	Students number
K M Shah Dental College	64
Manubhai dental college	36
Total	100

A total of 25 experts contributed their insights and expertise in the finalization of various constructs and items for the scale. Initially, 15 specialists were involved in the development phase before the pilot test, helping to refine a comprehensive 35-item questionnaire.

Following this,10 additional experts participated in the validation process, meticulously evaluating and confirming the relevance and efficacy of an updated 20-item questionnaire.

Phase- 2 sample

Implementation of scale with 410 postgraduate students

30 students participated separately for the test and retest reliability of the scale

The formula used for sample size determination is as per Slovin's formula, and available data obtained from the DCI site for dental postgraduates in India.

P= 6693, d =0.05, p= as per data available on DCI Year 2021-22

N = P/ (p x d2) + 1

= 6696/ (6696 x 0.05 x 0.05) + 1

= 6696/ (6696 x 0.0025) + 1

= 6696/17.74

= 377.45

Therefore, the minimum sample size required was 378 postgraduate students for the implementation of the scale. Out of 415 students who participated in the online mode with consent, the data from five students who didn't complete the questionnaire properly was excluded, leaving a total of 410 students for data evaluation.

Final sample -100 postgraduate students for a pilot study

410 for implementation of scale

30 students for test-retest reliability

3.9.1 Sampling method

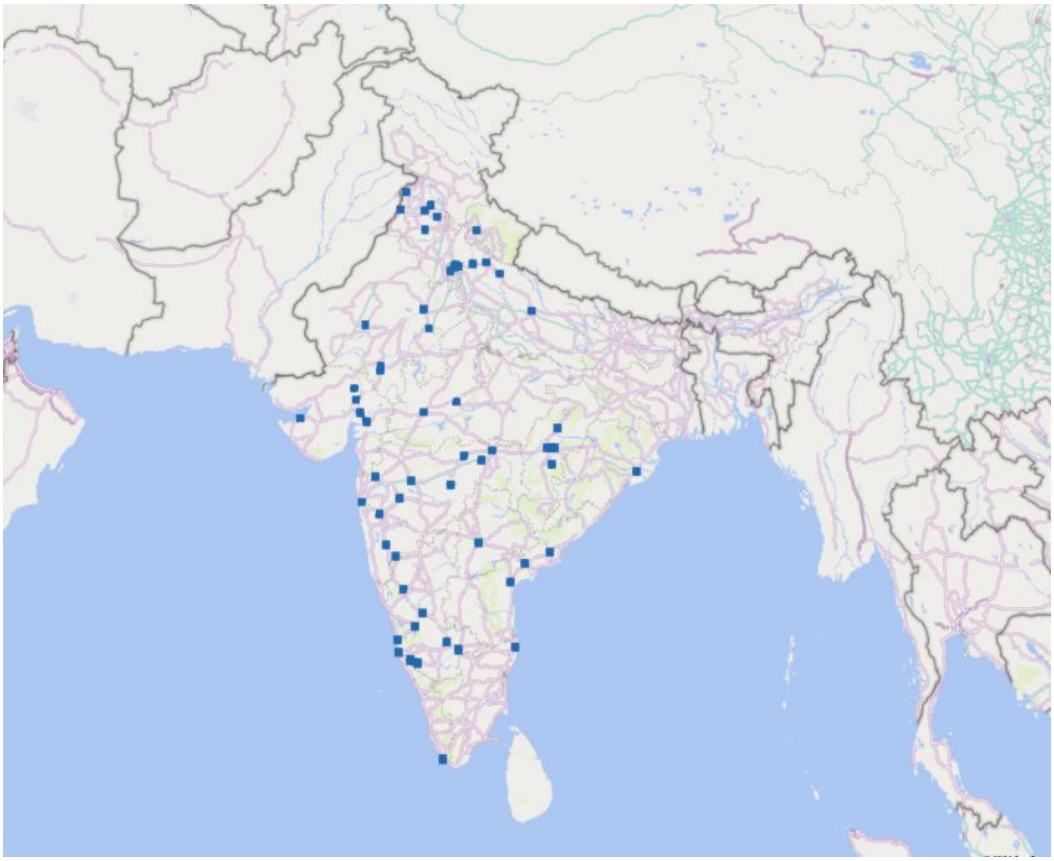
Random sampling method was utilized for the selection of colleges offering post graduate courses in India. The purposive sampling method was used for taking the postgraduate students.

3.9.2 Setting and Participants

After conducting a pilot study and finalizing a 20-question questionnaire, data was collected using an online survey created with Google Forms. Students were informed about the study's purpose and scope, and the link to the online survey was shared in various postgraduate WhatsApp groups available across India.

The questionnaire also included information about the purpose of the study, and participants were asked to give consent to the statement, "I consent to participate in the study." This permitted online consent from all the participants.

Fig 3.1 Representation of participating Dental colleges of various cities of India in Map



All the participating colleges are listed in Annexure 3

3.10 Tool

A research instrument has a crucial role in, acting as the foundation for accurate data assembly. It significantly affects the reliability of the findings, enabling researchers to draw meaningful conclusions that ultimately contribute to developing effective solutions for the issues being investigated.

Typically, a research instrument contains a series of carefully crafted items that have been carefully chosen and refined based on specific criteria to ensure their relevance and effectiveness. In the context of a professionalism scale, these items are referred to as statements. These statements can be understood as assertions about a psychological object; they encapsulate the essence of the concept being measured. Participants interact with these statements by indicating their level of agreement or disagreement and providing valuable perceptions about attitudes related to the subject.

3.10.1 Professionalism Scale

Initially, a total of 50 items were created after a comprehensive review of the literature. Following in-depth discussions and analysis, the researcher who has a personal expertise in dentistry and a deep interest in professionalism collaborated with two senior experts from the Departments of Education and Psychology to narrow the statements down to 35. An in-depth examination of the elements that constitute Professionalism reveals a noteworthy lack of consensus among researchers regarding a singular definition of the term. As a result, efforts have been made to enhance clarity while operationalizing the concept of Professionalism. However, it has been observed that crucial components such as sustainability, personal upliftment, lifelong learning, and work-life balance are frequently overlooked in many existing scales within the literature. Furthermore, no existing scale has successfully integrated all four essential domains: Ethical, Personal, Professional, and Societal. To address this gap, the researcher developed a comprehensive set of 35 statements, ensuring an equitable distribution of focus across the four domains, with 5 statements dedicated to Ethical aspects and 10 each to Personal, Professional, and Societal dimensions. This approach aimed to deliver a more holistic view of Professionalism.

3.10.2 Format and Nature of Statements

Participants assessed each statement on a five-point rating scale to express their level of agreement with various value statements. In this scale, a rating of 1 corresponded to "strongly disagree," while a rating of 5 indicated "strongly agree." The available response options were "strongly agree," "agree," "neutral," "disagree," and "strongly disagree," allowing participants to convey their opinions accurately. The original professionalism scale consisted of 35 items, thoughtfully crafted to encompass both positive and negative polarity statements, thereby providing a balanced perspective on the constructs being measured.

- Strongly disagree 1
- Disagree 2
- Neutral 3
- Agree 4
- Strongly Agree 5

Table 3.1 Distribution of statements for the professionalism scale used in the Pilot study

Components		Number of Statements with Positive Polarity	Number of Statements with Negative Polarity
Ethical	**5**	**4**	**1**
Personal	**10**	**8**	**2**
Professional	**10**	**8**	**2**
Societal	**10**	**8**	**2**
Total	**35**	**28**	**7**

Table: 3.2 Initial 35 statements scale

Statement Number	Statements	Domain	Abbreviation
1	I don't believe in excellence	Personal	P
2	I believe in teamwork and collaboration	Societal	Sc
3	I maintain good resiliency/adaptability in all circumstances	Personal	P
4	I believe in respecting the Autonomy of my patients	Ethical	E
5	I don't believe in job satisfaction	Professional	Pr
6	I believe in the prevention, promotion, and maintenance of public health	Societal	Sc
7	I relate my work with prestige	Professional	Pr
8	I don't believe in having a successful career	Professional	Pr
9	I am a good listener and communicate effectively with my patients	Professional	Pr
10	I believe in supporting people of all communities (solidarity)	Societal	Sc

11	I believe in respecting the Confidentiality of my patients	Ethical	E
12	I believe in maintaining good physical/mental and spiritual well-being.	Personal	P
13	I believe in sustainable living	Societal	SC
14	I don't believe in acting for the benefit of my patients	Ethical	E
15	I don't want to earn a good living	Personal	P
16	I want to promote healthcare leadership	Societal	Sc
17	I believe in using social media ethically	Societal	Sc
18	I believe in lifelong learning	Professional	Pr
19	I maintain a good work-life balance	Personal	P
20	I practice self-discipline	Personal	P
21	I don't believe in the evolution of society	Societal	Sc
22	I want to care for the public	Societal	Sc
23	I don't want to be accountable to my community	Societal	Sc
24	I am compassionate toward my patients	Professional	Pr
25	I want to be a perfectionist.	Personal	P
26	I practice self-reflection	Personal	P
27	I have good professional integrity	Professional	Pr
28	I enjoy a good quality of life	Personal	P
29	I am honest and maintain personal integrity	Personal	P
30	I believe in taking accountability for my work toward my patients	Professional	Pr
31	I believe in giving equitable justice to my patients	Ethical	E
32	I work passionately for my patients	Professional	Pr
33	I believe in doing no harm to my patients	Ethical	E
34	I am kind to my patients	Professional	Pr
35	I believe in community service	Societal	Sc

3.11 Experts validation: A total of 25 experts contributed their insights and expertise in the finalization of various constructs and items for the scale. Initially, 15 specialists were involved in the development phase before the pilot test, helping to refine a comprehensive 35-item questionnaire.

The validation process included participation from 25 dental practitioners and faculty members from various colleges across India. These participants comprised both academicians and private practitioners who graduated from dental school between 1965 and 2005. The average age of the experts was around 45 years, with a standard deviation of ±10 years. The group included 12 males and 13 females.

This exercise was performed to collect the insights of specialists on the highlighted components concerning their suitability, significance, and capacity to represent the concept of professionalism. A copy of the correspondence sent to the specialists and the roster of experts can be found in Appendices I and II, respectively.

The majority of the specialists who replied to the correspondence agreed that the identified components and statements accurately captured the core of professionalism, considering them both suitable and significant. However, several experts suggested merging some of the components due to their similarities, proposing that a more streamlined approach could improve clarity and focus.

Table 3.3: The responses of 15 experts for Initial 35-item questionnaires

Expert	**Suggestions**
1	Suggested to make the statements as simple as possible, and they should be direct and have positive meaning.
2	Include Questions related to personal components such as individual traits and work-life balance.
3	Suggested changes in Question number 2,7,11,34
4	Suggested to keep the Language simple.
5	No major suggestions suggested
6	Valued the efforts and suggested minor changes in questions 3,8 12 and 34
7	Suggested to use of the societal components in modest language.
8	Advised changes in questions related to quality of life
9	Suggested changes in question 3,12,15,20,34

10	Appreciated the scale
11	Suggested changes in question 11,15,34 and to Avoid the use of negative statements.
12	Suggested reframing Questions no. 2,5,6,11, 29, and 34 and to make changes.
13	Suggested to make simple statements with no negative aspects.
14	Suggested that the ethical component should be added in simple language.
15	Suggested that Statements and constructs with similar or overlapping meanings should be removed or merged

3.12 Pilot Study:

It was done for the final selection of statements A pilot study was done on 100 students with a scale with 35 questions. The data was interpreted as follows**.**

For the final selection of statements, the investigator followed the method proposed by Likert (1932), which involved analysing the differences between high and low groups. Initially, the investigator examined the frequency distribution of scores from all responses. They selected 27% of participants (NH = 27) with the highest total scores and 27% (NL = 27) with the lowest total scores for the item analysis.

These participants were categorized into high and low groups. To evaluate the responses from both groups for each statement, a T-value was calculated. The T-value indicates how well a statement distinguishes between the high and low groups. As a rough guideline, a T-value of 1.75 or higher suggests that the average responses of the high and low groups differ significantly, with a p-value greater than 0.05 also considered significant.

To conduct the T-test, random serial numbers were assigned to the 35 statements. From the table, it's clear that statements numbered 1, 5, 7, 8, 12, 15, 16, 21, 22, 23, 25, 28, 32, 34, and 35 did not meet the specified criteria. Consequently, these 15 statements were excluded, resulting in the final selection of 20 statements for the professional commitment scale.

Table 3.4: Mean, SD, T- value & P Value for Thirty-Five items of Professionalism Scale

Statement Number	Mean		SD		T - Value	P - Value
	Upper	Lower	Upper	Lower		
1	4.37	4.19	0.74	1.00	0.772	0.441
2	4.70	4.07	0.47	0.73	3.78	0.000
3	4.85	3.85	0.46	0.91	5.117	0.000
4	4.78	3.96	0.42	0.71	5.142	0.000
5	4.37	4.19	0.74	1.00	0.772	0.450
6	4.85	3.96	0.36	0.71	5.821	0.000
7	4.37	4.19	0.74	1.00	0.772	0.420
8	4.15	4.19	0.82	1.00	-0.149	0.885
9	5.00	4.00	0.00	0.55	9.367	0.000
10	5.00	3.93	0.00	0.68	8.266	0.000
11	5.00	3.93	0.00	0.38	14.5	0.000
12	4.15	4.19	0.82	1.00	-0.149	0.882
13	4.96	4.00	0.19	0.55	8.522	0.000
14	5.00	4.26	0.00	0.45	8.619	0.000
15	4.15	4.19	0.82	1.00	-0.149	0.879
16	4.26	4.15	0.71	0.99	0.474	0.631
17	4.37	3.19	0.88	0.83	5.069	0.000
18	4.85	3.67	0.36	0.55	9.297	0.000
19	4.74	3.59	0.45	0.64	7.677	0.000
20	4.74	3.56	0.45	0.70	7.432	0.000
21	4.26	4.15	0.71	0.99	0.474	0.629
22	4.37	4.19	0.74	1.00	0.772	0.444
23	4.33	4.19	0.73	1.00	0.62	0.535
24	5.00	4.04	0.00	0.34	14.822	0.000
25	4.26	4.15	0.71	0.99	0.474	0.638
26	5.00	4.11	0.00	0.42	10.902	0.000
27	4.96	4.00	0.19	0.55	8.522	0.000
28	4.33	4.19	0.73	1.00	0.62	0.513
29	5.00	4.04	0.00	0.71	7.086	0.000
30	4.93	3.85	0.27	0.46	10.562	0.000
31	4.37	3.26	0.88	0.90	4.571	0.000
32	4.33	4.19	0.73	1.00	0.62	0.529
33	4.74	3.44	0.45	0.75	7.709	0.000
34	4.33	4.19	0.73	1.00	0.62	0.538
35	4.26	4.15	0.71	0.99	0.474	0.640

For the Pilot study, the percentile norms were calculated by computing the mean, standard deviation, median, percentiles, skewness, and kurtosis for the total sample (N= 54).

Table 3.5: Percentile norms of Professionalism Scale (Pilot study)

Descriptive Statistics of Samples	
Mean	150.24
SD	16.88
Median	153
P10	128.5
P20	138
P30	140
P40	149
P50	153
P60	157
P70	160
P80	165
P90	171.5
Skewness	-0.914
Kurtosis	1.032

The minimal score recorded on the Professionalism Scale was 97, while the highest was 175. A total of 24 students (44.4%) exhibited low professionalism values, with scores at or below 152. The average professionalism score for 4 students (7.4%) was determined to be 153. The professionalism score of 26 students (48.1%) was identified as high, being equal to or exceeding 154.

Following a comprehensive pilot study, 15 statements were eliminated, (the procedure of which is given in chapter 4), leaving only 20 for further evaluation. A panel of 10 experts highlighted the critical importance of simplicity in the scale's design. They strongly recommended removing any negative statements to ensure a consistently positive framework throughout the assessment.

Every remaining statement was carefully rephrased to adopt a positive tone, aligning with the experts' suggestions for clarity and straightforwardness. As a result, the scale evolved into a purely positive format, leading to the decision to exclude all negative polarities. Consequently, this revised scale can no longer be classified as a traditional Likert scale; instead, it is best described as a five-point rating scale, designed to provide clearer, more affirmative feedback.

Following this, 10 additional experts participated in the validation process, meticulously evaluating and confirming the relevance and efficacy of an updated 20-item questionnaire obtained after the pilot study.

Table: 3.6 The responses of 10 experts for the 20-item questionnaire

Expert	**Suggestions**
1.	Valued the scale and suggested not to keep negative statements
2.	No changes suggested
3.	Appreciated the components
4.	Suggested to avoid the use of negative statements and make it simple and understandable scale.
5.	No change was suggested.
6.	Suggested to avoid the usage of negative statements as it makes the scale difficult to comprehend.
7.	No suggestions. liked the scale

8.	Suggested to use positive statements only
9.	Found the scale appropriate for the post-graduate dental students
10.	Appreciated the questionnaire

3.13 THEORETICAL FRAMEWORK OF STANDARDIZATION OF SCALE

Standardization creates uniformity and consistency in research measurements, procedures, and practices. It involves developing standardized tools and techniques that can be replicated across studies, confirming that the data collected is comparable and consistent. Standardization is essential in research because it allows for accurate data comparisons across studies, facilitating the identification of patterns and trends in a particular field of study. Standardization is achieved through a series of steps that involve designing, piloting, and refining research tools and procedures. These steps ensure that the research tools and methods are valid, reliable, and effective in measuring the variables under study. Standardization also involves developing clear and consistent data collection, management, and analysis guidelines, ensuring that research is conducted consistently and transparently. There are many benefits of standardization in research. Standardization ensures that the data collected is reliable, valid, and comparable across studies, making identifying patterns and trends in a particular field easier. It also reduces the potential for bias in research, ensuring that research is conducted fairly and accurately. Standardization also makes it easier to replicate studies, allowing for the validation of findings and establishing consensus in a particular field of study. Therefore, standardization is an essential component of research. It confirms that the information collected is reliable, valid, and comparable across studies, facilitating the identification of patterns and trends in a particular field of study. Standardization reduces the potential for bias in research, making research fairer and more accurate. It also makes it easier to replicate studies, allowing for the validation of findings and establishing consensus in a particular field of study.

Standardizing the scale is crucial to getting a reliable and valid assessment tool. The process involves establishing norms, procedures, and criteria for the scale's administration, scoring, and interpretation. The objective is to confirm that the scale evaluates the intended concept accurately and reliably and that the scores obtained on the scale are meaningful and interpretable.

The standardization process begins with selecting a characteristic sample of individuals similar to the cohort for which the scale is intended. The sample is typically chosen according to specific criteria, such as age, gender, education level, or occupation, to confirm that it is demonstrative of the population. The sample size is calculated based on statistical principles to certify that the results are dependable and valid.

Once the sample is selected, the scale is administered to the participants. The information obtained from the scale administration is evaluated to establish the scale's norms, procedures, and criteria. Norms refer to the average scores received on the scale by the sample and are used to compare the scores of individual participants to the cluster. Procedures refer to administering, scoring, and interpreting the scale. Criteria refer to the standards used to determine the meaning of the scores obtained on the scale.

The standardization process also involves assessing the reliability and validity of the scale. The consistency and uniformity of the scores obtained on the scale are referred to as reliability. The precision of the scale in measuring the intended concept is referred to as validity. Through the use of statistical techniques like factor analysis, correlation analysis, and item analysis, the validity and reliability of the scale are evaluated. It is a rigorous and complex process that ensures that the scale is reliable, valid, and fair and that the scores obtained on the scale are meaningful and interpretable. It involves selecting a typical sample, administering the scale, analyzing the data, creating the norms, procedures, and criteria, and assessing the reliability and validity of the scale. Standardizing scale methodology is essential to developing a valuable and accurate assessment tool.

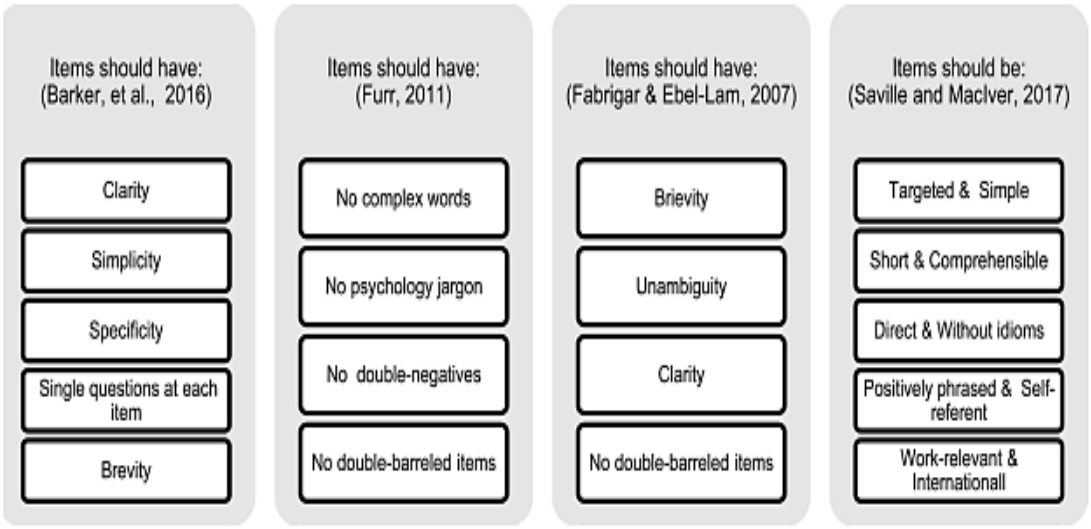


Fig 3.2

Key principles for successful item writing (Streiner et al., 2015)

A questionnaire is a group of items planned to evaluate one or more primary constructs, often known as latent variables. It consists of standardized, objective self-report questions, which are added up to provide a score. . Scales, such as ability, attribute, and attitude scales, serve as markers of the measured construct, and their scores serve as an indication of the construct.

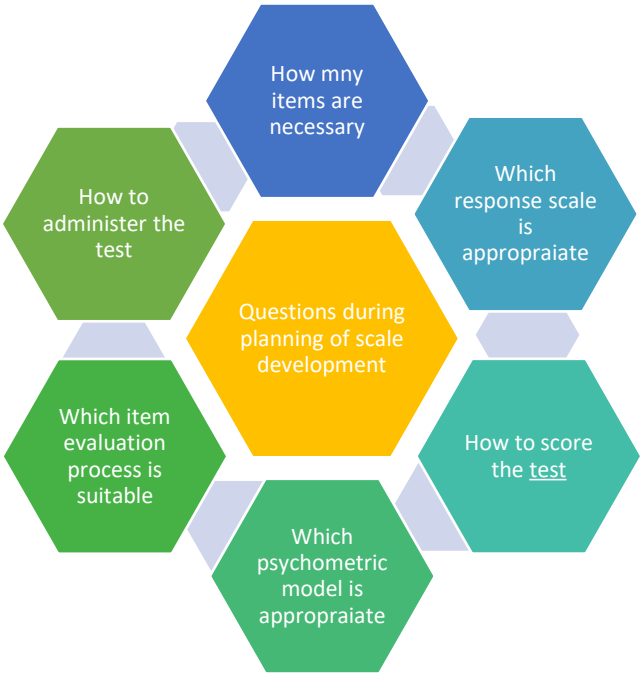


Fig 3.3 Planning of Scale Development (Irwing & Hughes, 2018)

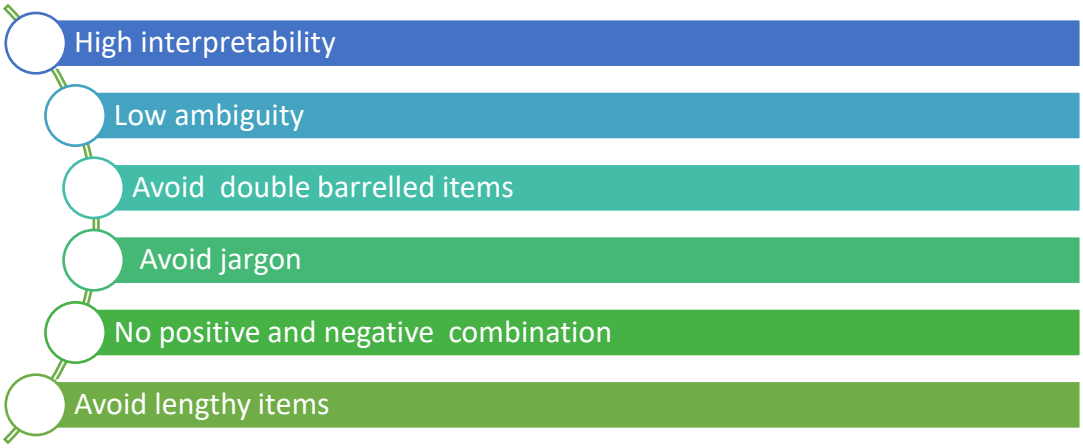


Figure 3.4 **Proposed Criteria for retaining and discarding items before or after expert reviewing (Streiner et al., 2015)**

Scale creation, also known as scale development, is the process of selecting and/or creating the best possible test items for a specific population. Standardization, or norming, is a process that comes after test development in the related field of test development. The final pool is created during test creation once the items that are the best show the latent construct being measured have been assembled and analyzed. An original scale creation has a substantial impact on research findings. It also affects the instruments' sensitivity and accuracy.

Trochim (2006) outlines the scale development process as a series of steps. First, assume the attribute is unidimensional and define it. Then, create a pool of 80–100 Likert items and evaluate them on a 5 or 7-point scale. Next, have a panel of experts score the items, and decide which ones to keep for the final scale.

Finally, apply the scale and reverse any items that measure something in the opposite direction from the rest. The measurement accuracy of the total score is crucial as the respondent's scores on all questions determine the overall valuation with an instrument.

An "item" is a single question or statement that participants are supposed to react to or answer. Because of its uniform answer structure, a set of items intended to evaluate a common notion is referred to as a "scale." An alternative "survey" can consist of a set of scales or random questions with no clear-cut right or incorrect answers. It might also include different response styles for the scales and/or questions.

Content validity is a crucial aspect of research, assessing how well an instrument accurately represents the intended content domain. It is typically assessed through expert judgment and

review of the instrument's content. Content validity is a criterion validity that ensures a study accurately captures the concept it is supposed to represent. Researchers must verify the validity of the content through methods such as focus groups, polls, and expert opinion. Content validity is essential for any outcome measure and ensures that the measure provides a complete and accurate picture of the most important parts of a concept or attribute for a known audience and setting.

Formula: The content validity ratio (CVR)

$$\mathrm{CVR} = \frac{n_e - \frac{N}{2}}{\frac{N}{2}}$$

Construct validity is the evaluation of a measurement tool's accuracy in measuring the intended construct. It is based on factor analysis and correlation analysis to determine if the instrument produces scores consistent with the measured theoretical construct. Construct validity is more challenging to demonstrate than other forms of psychological validity due to the abstract nature of constructs. It considers empirical data supporting the test's ability to measure the construct and the theoretical association between the test and that construct. Other types of evidence, such as criteria, content, discriminant, and convergent validity, also support construct validity. Construct validation is the continuous learning process about a test's measures, the significance of test results, and the suitability of interpretations and applications of test results.

Reliability is the consistency and stability of a measuring tool over time and with different samples. It assesses the instrument's measurements and should produce results when given to the same individuals or those assumed to have the same characteristics. Reliability is typically evaluated using statistical techniques like test-retest reliability and internal consistency reliability. A reliable method should consistently produce similar outcomes when the same procedure is used on the same sample in the same circumstances, preventing potential bias and questionable measurement techniques. Reliability is a critical factor in research, ensuring that

the measurements and results obtained are consistent and accurate. When conducting investigations, researchers take into account many levels of reliability.

Test-retest reliability is one type that examines the consistency of a test or measure over time. Researchers give the same test to the same respondents on multiple dates, compare the results, and determine the test-retest reliability.

Internal consistency and reliability refer to the degree to which different items or questions in an exam produce consistent results. For example, if a survey asks questions about satisfaction with various aspects of a service, the responses should be consistent and not contradict each other. The Cronbach's alpha coefficient can be utilized by researchers to judge dependability and internal consistency.

Exploratory factor analysis (EFA) is a statistical technique used to determine the primary structure of a set of variables. It seeks to identify the fundamental connections between the variables that are observed by examining their correlations and grouping them into factors that explain the most variation in the data. Factors represent the basic constructs or dimensions that the pragmatic variables measure. EFA differs from confirmatory factor analysis (CFA), which tests a specific theoretical model of the relationships between variables

In 1904, psychologist and statistician Charles Spearman invented factor analysis for his intelligence research. The strategy has since evolved and comes in two primary flavors: confirmatory and exploratory. In EFA, every observed variable may be a measure of every factor, to identify the strongest connections between factors and observed variables.

EFA has many applications in various fields, including psychology, sociology, education, and marketing research. It can be used to develop and validate measurement scales, identify factors that influence a particular behavior or attitude, and reduce the complexity of large data sets. Researchers should use EFA when they have no prior preconceptions about the components or patterns of the measured variables.

The foundation of Exploratory Factor Analysis (EFA) lies in the concept of common factors,

which suggests that any measurable variable can be related to multiple aspects. Before employing Confirmatory Factor Analysis (CFA) to develop a scale, researchers should first conduct EFA. EFA is essential for identifying the underlying structures or factors in a set of measured variables, while CFA allows researchers to test whether there is a relationship between the assessed variables and their underlying latent factors or constructs.

The KMO test calculates the partial correlation coefficient or the degree to which the factors clarify each other, between the variables. When KMO levels are closer to 1.0, they are marked great; when they are less than 0.5, they are considered undesirable. Nowadays, the majority of scholars argue that factor analysis can start with a KMO of 0.80 or above.

Test of Bartlett's Sphericity Bartlett's test for sphericity assesses the null hypothesis that the correlation matrix is an identity matrix. An identity correlation matrix indicates that your variables are independent and unsuitable for factor analysis. When a significant statistical result, typically with a significance level below 0.05, indicates that the correlation matrix deviates from being an identity matrix, the null hypothesis is dismissed.

Descriptive statistics is a fundamental tool for comprehending and communicating information through data collection, organization, analysis, and interpretation. Its primary goal is to provide a concise and understandable summary of the data for researchers, analysts, and decision-makers.

Central tendency metrics like mean, median, and mode are frequently employed. A set of data's mean is determined by adding up all of the values and dividing the total by the number of observations. In a sorted data set, the mode is the value that is observed most frequently, while the median is the middle-found value.

Variability measures, including range, variance, and standard deviation, are used to exemplify the spread or dispersion of the data. The change between a data set's maximum and minimum values is identified as the range. The standard deviation is the square root of the variance, which is determined by calculating the average squared difference between each value and the mean. These measures provide insights into how widely dispersed the data is from the central point.

Descriptive statistics also consider the distribution of the data, which can be symmetrical, skewed, or bimodal. Skewed distributions can be either negatively or favorably skewed (longer tail on the left or right). Graphical representations, such as

Histograms, bar charts, line graphs, and scatter plots, are commonly used to present descriptive statistics. These visual aids help identify patterns, trends, and outliers in the data. In inference, descriptive statistics are crucial for briefing and unfolding data accurately and effectively. They provide a clear and brief data summary that can be easily interpreted and understood by the audience. Descriptive statistics are widely used in numerous fields like science, economics, medicine, and the social sciences.

The implementation of scale in methodology is an imperative step in certifying that a scale measures the intended construct accurately and consistently. The process includes establishing norms, procedures, and criteria for administering, scoring, and interpreting the scale effectively. To implement scale methodology effectively, researchers need to follow a systematic and rigorous approach, which typically involves the following steps:

1. Pilot testing: Before implementing the scale, researchers should pilot it with a small sample of individuals to recognize any issues or problems with the scale's administration, scoring, or interpretation. Researchers can make necessary adjustments to the scale based on the pilot test results

2. Sampling: Researchers should select a characteristic sample of individuals similar to those for which the scale is intended. The sample size must be calculated according to statistical principles to ensure that it is large enough to yield reliable and valid results.

3 Standardization: Researchers should develop explicit norms, procedures, and criteria for administering, scoring, and interpreting the scale. This includes defining the scale's response options and scoring system, training assessors, and establishing guidelines for handling missing data.

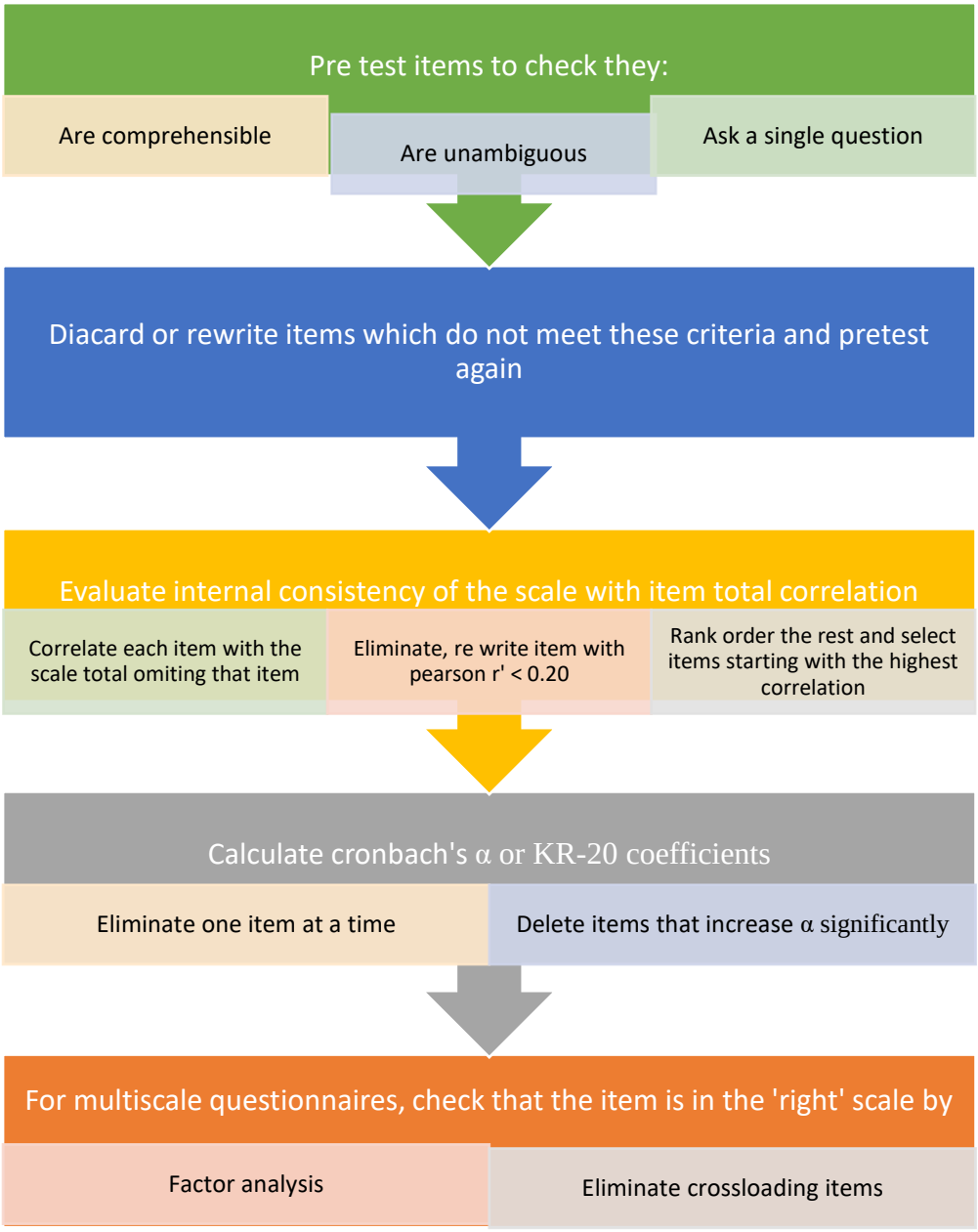


Figure 3.5 Overview of Pilot Testing Procedure Streiner et al. (2015: p. 94).

4 Data collection: Investigators should gather data from the sample using the standardized procedures and criteria established in the previous step.

5 Data analysis: Researchers should analyze the data collected using appropriate statistical techniques to assess the reliability and validity of the scale. This requires analyzing the uniformity within the scale itself, its ability to produce consistent results over time, and how well it assesses what it is designed to assess.

6 Interpretation: Researchers should interpret the outcomes of the data analysis to describe the scale's usefulness and applicability in measuring the construct of interest. This includes examining the scale's sensitivity, specificity, and predictive validity.

Figure 3.6 Phases of Scale Development (Kyriazos& Stalikas, 2018)