Instrument used to assess interprofessional education and collaborative practice in health professional students: A COSMIN systematic and psychometric review

Anastasia Anna*1 , **Ying-Chin Liao2** , **Linlin Lindayani3** , **Aan Nuraeni1**

1 Faculty of Nursing, Universitas Padjadjaran, Bandung, West Java, Indonesia
2 Department of Nursing, Dalin Tzu Chi Hospital, Chiayi, Taiwan
3 Sekolah Tinggi Ilmu keperawatan PPNI Jawa Barat, Indonesia

Abstract

Background: Interprofessional education and collaborative practice (IPE/IPC) are essential for preparing students to work together and respect the unique qualities and abilities of professionals. However, IPE/IPC and its related concepts are highly abstract phenomena and complicated to assess and measure. In consequence, a critical appraisal is needed to evaluate the quality of the instruments.

Purpose: This study aimed to critically appraise, compare and summarize the quality of measurement properties of all self-report collaboration questionnaires for health professional students and to provide evidence concerning the psychometric properties of the measurement.

Methods: A psychometric review was employed, and the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) approach was applied to assess the methodological quality of the nature of the measurements. Data search using keywords: health professional students, interprofessional, collaboration, teamwork, collaborative, through MEDLINE, EMBASE, and EBSCO-hosted Education Resource Information Centre databases.

Results: Seven instruments from 10 reviewed studies were identified. Among them, four instruments targeted attitudes toward collaboration. One instrument focused on students' collaborative learning readiness and had been tested in Hong Kong using English, in Iran using Persian, and in Indonesia using Bahasa Indonesia. One instrument measured perception about IPE, and two studies measured IPE/IPC competencies related to patient-centered, team-based, and collaborative care. The methodological quality assessment indicated that several instruments were less rigorously developed and validated than COSMIN and Quality Criteria of Measurement Properties recommend.

Conclusion: The findings of this psychometric review are that the Interprofessional Attitudes Scale is an instrument with adequate content validation and very good structural validity, internal consistency, cross-cultural validity, reliability, measurement error, and criterion validity. It is recommended that the Interprofessional Attitudes Scale be used to measure the interprofessional attitudes of health professional students.

Keywords: collaborative practice; interprofessional education; instrumen;, psychometric review; validity

Introduction

The process of establishing and sustaining productive interprofessional working relationships between healthcare students and professionals, patients and families, and communities to promote optimal health outcomes is referred to as interprofessional collaborative practice (IPC) (World Health

Organization [WHO], 2010). Integrating IPC and coordinated health professional activities can result in optimal patient outcomes, which may be achieved through the implementation of interprofessional education (IPE) on an organized, systematic basis (Hojat & Herman, 1985). It is, therefore, crucial to equip students with the necessary skills and knowledge from an early stage of their education to enable them to work effectively in a collaborative environment, recognizing the value of diverse professional expertise in the sharing of information, problem-solving and the completion of learning activities and final projects (Schmitz et al., 2017; Tibi, 2015).

Consequently, IPE and IPC competencies should be assessed starting from school education and continuing throughout the nursing career (Kajander-Unkuri et al., 2014). However, both IPE and IPC are very abstract and complex phenomena to assess and measure (Hojat et al., 2014). Clear and unambiguous operational definitions of the

concepts are essential to guide the development of valid IPE and IPC assessment instruments (Hojat et al., 2014). Although numerous concept analyses have been conducted to define IPE and IPC, there is limited consensus on their definitions and associated measures (Dominguez et al., 2015).

Nevertheless, a variety of instruments have been developed to measure IPE and IPC outcomes. However, the lack of rigorous evaluation of the effectiveness of IPE curricula makes it challenging for educators to ensure that students are adequately prepared to work with collaborative practices. Furthermore, in order to provide credible data for the testing of IPE effectiveness, instruments must be developed with strong psychometric properties, including validity and reliability. Appropriate scale and score construction permit the differentiation of test takers and facilitate the interpretation of test scores in a valid manner (Oates & Davidson, 2015). Furthermore, high methodological quality and a thorough systematic review will produce the
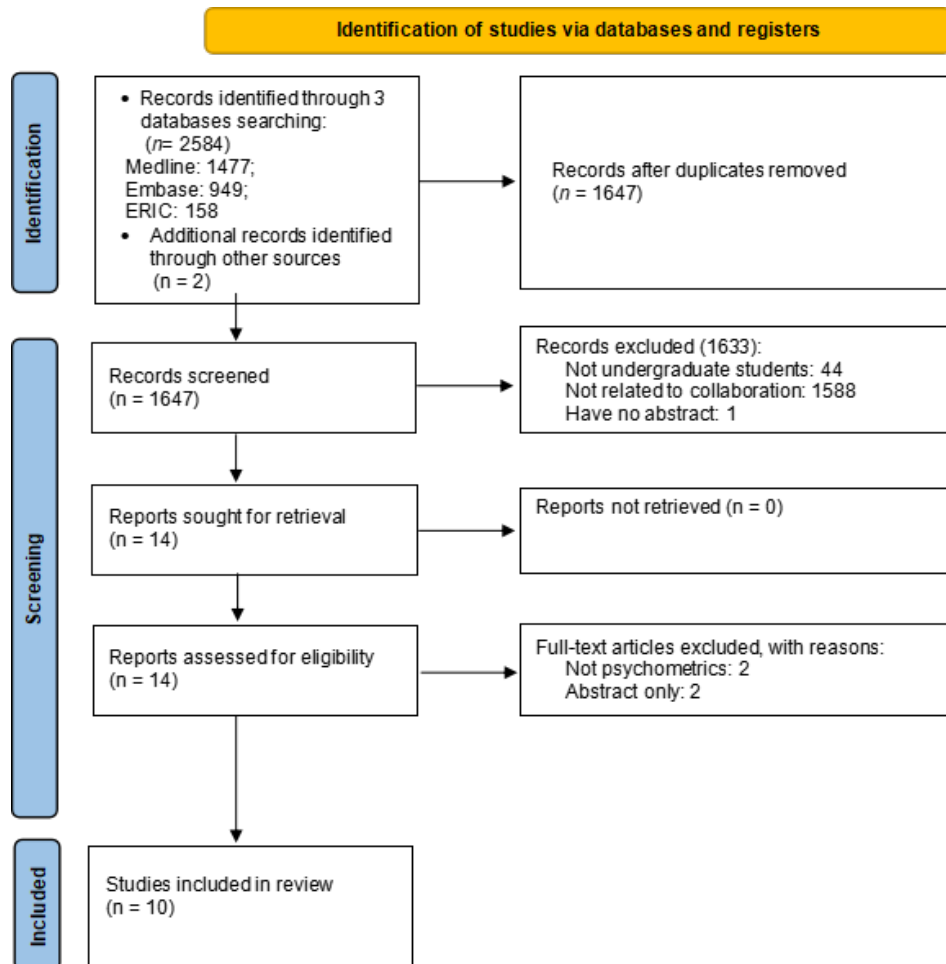

Figure 1.The PRISMA flowchart of studies selection

most appropriate measurement tool, one of which is by conducting a critical assessment to assess the quality of the instrument (Mokkink et al., 2010).

Oates and Davidson (2015) state that to critically assess IPE instruments, the Quality Assessment Scale for Interprofessional Learning (QuAILS) can be used, which is a standard checklist explicitly developed for review. Nine instruments in QuAILS measure attitudes toward collaborative learning, student perceptions, and student readiness for collaborative learning. In addition, the instrument also assesses the interprofessional socialization process and key aspects of interprofessional. Therefore, QuAILS meets the criteria set by the Educational and Psychological Testing standards for the design and development of educational tests (Oates & Davidson, 2015).

In contrast, the authors identified several areas that need to be improved, such as information on item design quality criteria and types of validity evidence that cannot be assessed using QuAILS. Therefore, it is important to "repeat" the review for IPE/IPC instruments, especially using the COSMIN framework. The objective of this study was to undertake a critical assessment, comparison, and synthesis of the quality of measurement properties of all self-report outcomes of interprofessional education (IPE) or interprofessional collaboration (IPC) measurements for health professional students. Furthermore, the study aimed to provide evidence regarding the psychometric properties of the measurements using the COnsensus-based Standards for the Selection of health Measurement INstruments (COSMIN) approach.

Methods

Design
A psychometric review was employed by applying the COSMIN approach to assess the study's methodological quality on the nature of the (Mokkink et al., 2010; Terwee et al., 2012). Firstly, the formulation of the research purpose, selection of databases and keywords, specification of inclusion and exclusion criteria, identification of databases to be searched, and the selection and extraction of data were guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). (Moher et al., 2009). Next, the psychometric properties of the instruments were qualitatively assessed using the Quality Criteria for Measurement Properties (Terwee et al., 2012). Then, data were analyzed and synthesized, and important findings were reported with summaries, charts, and figures.

Search methods
Two researchers (AA & YC) independently searched for articles published between 2000 and 2023 and met the inclusion criteria set together. The article search used the keywords health professions students, interprofessional, collaboration, teamwork,

and collaborative through the MEDLINE, Embase, and ERIC databases. AA & YC also carried out a manual search of the references in the included studies and retrieved and organized the hits from all the searches using RefWork®.

Inclusion and exclusion criteria
Inclusion criteria for this review included empirical research that (i) involved health professional students, (ii) focused on collaboration, interprofessional education, and collaborative practice, (iii) were all self-reporting, and (iv) assessed measurement properties.

Selection process
The literature search identified 2584 articles and two potential articles from the reference list. The first and third authors independently screened each publication's title and abstract, assessed full texts for eligibility, and, finally, decided on inclusion. The two reviewers (AA and LL) reached a consensus through discussion. Duplicates were identified and removed by AA and LL through a manual search of the databases searched in RefWorks© and between databases. After removing duplicates, the authors excluded 1633 articles after reading the titles and abstracts. Of these, 44 articles did not measure undergraduate students, 1588 articles were unrelated to interprofessional education and collaborative practice, and one article lacked an abstract. Two independent reviewers (AA and LL) assessed fourteen full-text articles for eligibility, and four articles were excluded because two of the studies did not deal with psychometrics. Two were abstracts only. Finally, the authors included ten studies in the qualitative synthesis. Figure 1 shows a PRISMA flowchart of the search process.

Data extraction
Two authors (AA and AN) performed data extraction for all included studies and resolved the differences through discussion between the authors. Seven different tools were found in the ten included studies. The study is characterised by the following features: the instrument used, the construct under investigation, the target population, the period of recall, the number of items included in each subscale, and the response options available (Table 1). After that, AA and AN also extracted the population's characteristics, including instrument, sample size, age, gender, psychometric administration (country, native language), response rate, and available translation (Table 2).

Quality appraisal
In order to ascertain the type of measurement property examined and to evaluate the methodological quality of the studies, the researchers employed the COSMIN checklist, which employs a 4-point scale (Mokkink et al., 2018). COSMIN is a methodology to critically appraise and assess the content validity for interprofessional education and collaborative

Anna, A., et al. (2024)

Table 1. Meaning of hope in children

Instrument	Construct	Target population	Recall period	Subscale(s) /(no.of items)	Respons options
The Scale of Attitude toward Physician-Pharmacist Collaboration (SATP2C) (Hojat et al.2012)	Attitudes toward collaborative relationships	263 3rd year students at medical college	N/A	16 items □Responsibility and accountability □Shared authority Interdisiplinary education.	4-point Likert scale. (1=strongly disagree, 4=strongly agree)
The student Perceptions of Interprofessional Clinical Education (SPICE-R)	Perceptions regarding IPE	277 1st year students in a variety of health profession-al degree programs	N/A	SPICE-R: 10 items □Outcomes □Teamwork □Roles responsibilities ATHCTS-R: 21 items □Team value □Team efficiency Shared leadership	5-point Likert scale. (1=strongly disagree, 5=strongly agree)
(Dominguez et al.,2014)	Attitudes toward interprofessional healthcare teams	490 students from some programs	N/A	14 items The quality of care (11 items). The time constraints (3 items)	5-point Likert scale. (1=strongly disagree, 5=strongly agree)
ATHCTS-R	The degree of read-iness of students to engage in interpro-fessional learning activities.	469 students in two univer-sities	N/A	18 items (item 18 was deleted) □Teamwork and collaboration □Positive Professional Identity □Negative professional identity Roles and responsibilities	5-point Likert scale. (1=strongly disagree, 5=strongly agree)
		200 final-year students	N/A	19 items □Teamwork and collaboration (9 items) □Negative Professional identity (3 items) □Positive Professional identity (4 items) □Roles and responsibilities (3 items)	5-point Likert scale. (1=strongly disagree, 5=strongly agree)
		800 1st, 2nd, and 3rd year students.	N/A	18 items (item 17 was omitted) □Teamwork and collaboration □Professional identity □Roles and responsibilities	5-point Likert scale. (1=strongly disagree, 5=strongly agree)

Cont. Table 1. Meaning of hope in children

Instrument	Construct	Target population	Recall period	Subscale(s) /(no.of items)	Respons options
The Interprofessional Attitude Scale (IPAS) (Noris et al., 2015)	The attitude related to core competencies of Interprofessional Collaborative Practice (IPEC)	1.549 students	N/A	27 items ▢Teamwork, Roles, and responsibilities (TRR) ▢Patient-centered (PC) ▢Diversity and ethics (DE) ▢Interprofessional Biases (IB) Community-centeredness (CC)	5-point Likert scale. (1=strongly disagree, 5=strongly agree)
The Jefferson Scale of Attitudes Toward Interprofessional Collaboration (JeffSATIC) (Hojat et al., 2014).	Attitudes toward interprofessional collaboration.	1976 health profession students in three universities: Philadelphia, USA. Chicago, USA. Victoria, Australia.		20 items ▢working relationships accountability.	7-point Likert scale. (1=strongly disagree, 7=strongly agree)
The International Collaborative Competency Attainment Survey (ICCAS) (Archibald, et al.,2014; Schmitz, et al.,2017)	The new IPE/IPC competencies	584 students from 15 interprofessional education programs in Canada and New Zealand.	Pre- and post-program.	20 items ▢Communication ▢Collaboration ▢Roles and responsibilities ▢A collaborative patient/family-centered approach ▢Conflict management/resolution. ▢Team functioning.	7-point Likert scale. (1=strongly disagree, 2= moderately disagree, 3= slightly disagree, 4 = neutral, 5= slightly agree, 6= moderately agree, 7=strongly agree; na=not applicable)
	Behaviors associated with patient-centered, team-based, collaborative care.	1.023 health profession students from six schools at the Academic Health Center.	Preand post-learning activities.	20 items ▢Communication ▢Collaboration ▢Roles and responsibilities ▢A collaborative patient/family-centered approach ▢Conflict management/resolution. ▢Team functioning.	5-point Likert scale. (1=poor, 2= fair, 3= good, 4 = very good, 5= excellent). Addition: 5-point Likert scale. (1=much better now, 2=somewhat better now, 3=about the same, 4=somewhat worse now, 5= much worse now). + interviews.

Anna, A., et al. (2024)

Table 2. Characteristics of Population

Instrument	Population				Instrument administration		Response rate	Available translation
	n	Age means (SD, Range)	Gender (% Female)	Characteristics	Country	Original language		
SATP2C	210	N/A	50	Medical students after the third-year clinical rotations	USA	English	80%	English
SPICE-R ATHCTS-R	221	Mean 24.4 SD 4.43	68	Nursing, optometry, pharmacy, physical therapy, and health administration students	USA	English	80%	English
ATHCTS	288	Mean 30.3 SD 9.8	81.65	Students from the law, nursing, medicine, pharmacy, social work, and dental schools.	USA	English	4.5%	English
RIPLS	469	Mean 20.95 SD 1.26	55.9	Two Hongkong university students from 5 healthcare professions: biomedical sciences, Chinese medicine, medicine, nursing, and pharmacy.	Hong-kong	Chinese	N/A	English
	200	N/A	N/A	Medical students	Iran	Persian	N/A	Persian
	755	N/A	73.8	Medicine, nursing, pharmacy, and public health students	Indonesia	Bahasa Indonesia	94.4%	Bahasa Indonesia
IPAS	678	N/A (%)	60.3	Students of the college of health: nutrition & physical therapy, college of nursing: undergraduate &graduate nursing, college of pharmacy, school of medicine: medical laboratory science, medical doctor program, physician assistant program, public health program.	USA	English	45%	English
JeffSATIC	1976	Mean 24.6 SD 4.3 Mean 24.0 SD 2.9 Mean 23.6 SD 6.2	63 50 63	Students from Jefferson Univ: Medical, nursing, pharmacy, occupational therapy, and physical therapy. Students from Midwestern: Pharmacy, medical. Medical science, and dentistry. Students from Monash: paramedic, occupational therapy, nutrition, physical therapy, and nursing.	USA	English	N/A	English
ICCAS	584	N/A	N/A	Undergraduate, postgraduate, and continuing professional development students from medicine, paramedicine, police, early childhood education, library, PT/OT assistant, support work, social work, nursing, massage therapy, dental hygiene, pharmacy, dietetics, kinesiology, spiritual care, dentistry, physiotherapy, occupational therapy, and health sciences.	Canada	English French	N/A	English
	785	Mean 21.9-26.4 SD 2.8-3.3	66.1	Health profession students from medicine, pharmacy, nursing, dentistry, veterinarian medicine, public health, occupational therapy, clinical laboratory science, social work, dietetics, and counseling psychology.	USA	English	77%	English

Table 3a. Result of Studies on Measurement Properties

Instrument	Country	n	Structural validity		Internal consistency (Cronbach's.α)		Cross-cultural validity/measurement invariance		Reliability (ICC/ Weight Kappa)		Pooled or summary result (overall rating)
			Rating	Result	Rating	Result	Rating	Result	Rating	Result	
SATP2C	USA	210	(?)	Not reported	(+)	0.90	(+)	No statistical difference between men & women (p > .05)	(?)	ICC= 0.82, 0.78, 0.68 Weight Kappa not reported	2(+), 2(?)
SPICE-R ATHCTS-R	USA	221	(+)	CFI 0.95, RMSEA 0.077	(+)	0.86	(?)	No multiple group performed	(+)	Weight Kappa ≥ 0.70	3(+), 1(?)
ATHCTS	USA	288	(+)	CFI 0.95, TLI 0.94, RMSEA 0.067, SRMR 0.04	(+)	0.92	(+)	p < .05 p < .01	(+)	ICC = 0.86	4(+)
RIPLS	Hong-kong	469	(+)	CFI 0.94, TLI 0.93, RMSEA 0.062	(+)	0.71 - 0.91	(?)	No multiple group performed	(?)	Weight Kappa not reported	2(+), 2(?)
	Iran	200	(+)	CFI 0.97, RMSEA 0.06	(+)	0.89, 0.60, 0.86, and 0.28	(?)	No multiple group performed	(?)	Weight Kappa not reported	2(+), 2(?)
	Indone-sia	755	(+)	CFI 0.94, TLI 0.93, RMSEA 0.063, SRMR 0.043	(+)	0.71	(+)	No important differences between group factors (age & language)	(+)	ICC = 0.87	4(+)
IPAS	USA	678	(+)	RMSEA 0.062.	(+)	0.61 - 0.92	(+)	No statistically significant differences	(?)	Weight Kappa not reported	3(+), 1(?)
JeffSATIC	USA	1976	(+)	IRT: χ2 > 0.01	(+)	0.84 - 0.90	(-)	the differences were statistically no significant (p < .05)	(?)	Weight Kappa not reported	2(+), 1(-), 1(?)
ICCAS	Canada	584	(?)	No info	(+)	0.98. 0.94, and 0.96	(?)	No multiple group performed	(?)	Weight Kappa not reported	1(+), 3(?)
	USA	785	(?)	No info	(+)	0.96	(?)	No multiple groups performed	(?)	Weight Kappa not reported	1(+), 3(?)

practice instruments in health professional students available in 2018 (Mokkink et al., 2018). AA and YC evaluated the sufficiency and quality of the measurement properties separately, encompassing nine distinct domains. These included content validity, structural validity, internal consistency, cross-cultural validity, reliability, measurement error, criterion validity, hypothesis testing, and responsiveness (Table 3a & Table 3b).

Firstly, two authors (AA & YC) independently appraised the ten boxes of risk of bias with the criteria as very good, adequate, doubtful, and inadequate. All changes were discussed, and researchers reached a full agreement. After that, LL & AN summarized them as sufficient (+), insufficient (-), or indeterminate (?). Then, all researchers (AA. YC, LL, and AN) discussed grading the quality of each instrument using the GRADE (Grading of Recommendations Assessment, Development, and Evaluation) approach for systematic review, and we graded the quality of evidence as substantial, moderate, low, or very low evidence.

Interpretability

Given the heterogeneity of the data, the primary author (AA) presented a descriptive and narrative synthesis of the data and subsequently tabulated the COSMIN quality ratings per study (Tables 4a and 4b) to illustrate the methodological quality of each study on measurement properties.

The COSMIN checklist comprises nine boxes (A–J), each containing methodological criteria for assessing a specific measurement property. The measurement properties assessed were as follows: A) internal consistency, B) reliability (test-retest, inter-rater, and intra-rater), C) measurement error, D) content validity, E) structural validity, F) hypotheses testing, G) cross-cultural validity, H) criterion validity and I) responsiveness.

Each box comprises five to 18 items (Mokkink et al., 2010). Each item was scored on a 4-point Likert scale (excellent, good, fair, poor). A methodological quality score was obtained for each measurement property based on the lowest rating of any box ('worst-score counts') (Terwee et al., 2012).

Results

From the ten (out of 3578) studies, we obtained seven IPE/IPC-related instruments: one instrument measured collaboration among interprofessional healthcare students, four instruments measured attitude toward collaboration (Hojat et al., 2012; Kim & Ko, 2014; Norris et al., 2015), one measured perception of interprofessional clinical education (Dominguez et al., 2015), three studies have reported on the development of an instrument designed to assess students' preparedness for interprofessional learning (Ataollahi et al., 2019; Ganotice & Chan, 2018; Tyastuti et al., 2014), and an instrument measuring collaborative competence was reported in two studies. (Archibald et al., 2014; Schmitz et al., 2017).

Measures of the Attitudes Toward Collaborative Relationship

The Scale of Attitude toward Physician-Pharmacist Collaboration (SATP2C) was developed in 2011. It is a four-point Likert scale, ranging from 1 (strongly disagree) to 4 (strongly agree). The authors developed 16 items, which were grouped into three subscales: responsibility and accountability, shared authority, and interdisciplinary education. A higher score indicates a more positive attitude toward the collaborative relationship between physicians and pharmacists (Hojat et al., 2012).

Furthermore, to gauge attitudes towards the collaboration of physicians and nurses, it is possible to utilize the scale developed for this purpose by Hojat et al. (2012), the Jefferson Scale of Attitudes towards Interprofessional Collaboration (JeffSATIC). The scale comprises 15 items answered on a four-point Likert scale (1 = strongly disagree, 4 = strongly agree). In 2014, Hojat et al. enhanced the scale, expanding it to 20 items and subdividing it into two subscales: working relationships and accountability (Hojat et al., 2014). They measured 1976 health profession students at two universities in the USA (Philadelphia and Chicago) and one in Australia (Victoria). The students scored on a seven-point Likert scale. It was found that the higher the scores, the more positive the attitudes towards interprofessional health students. So, this scale measures more than just the attitudes of collaboration between two professional students, likely SATP2C, but more professions.

In contrast, the original scale of the Attitudes toward Health Care Teams (ATHCTS) was developed in 1999 to measure the general attitudes of geriatric healthcare team

Table 3b. Result of Studies on Measurement Properties

Instrument	Country	n	Structural validity		Internal consistency (Cronbach's α)		Cross-cultural validity/measurement invariance		Reliability (ICC/ Weight Kappa)		Pooled or summary result (overall rating)
			Rating	Result	Rating	Result	Rating	Result	Rating	Result	
SATP2C	USA	210	(?)	MIC not defined	(+)	r=0.70	(+)	the result in line with 2 Hypo's (2+)	(?)	No hypothesis defined	2(+), 2(?)
SPICE-R ATHCTS-R	USA	221	(?)	MIC (-)	(?)	Not all info	(+)	the result in line with 3 Hypo's (3+)	(?)	No hypothesis defined	1(+), 3(?)
ATHCTS	USA	288	(?)	MIC (-)	(-)	r=0.38	(+)	the result in line with 2 Hypo's (2+)	(?)	No hypothesis defined	1(+), 2(?), 1(-)
RIPLS	Hong-kong	469	(?)	MIC (-)	(?)	r=0.13-0.68 Not all info	(+)	the result in line with 3 Hypo's (3+), Result not in line with 1 hypo's (1-)	(?)	No hypothesis defined Comparison between sub-group	1(+), 3(?)
	Iran	200	(?)	MIC (-)	(?)	Not all info	(+)	the result in line with 2 Hypo's (2+), Result not in line with 2 hypo's (2-)	(?)	No hypothesis defined. Comparison between sub-group	1(+), 3(?)
	Indonesia	755	(?)	MIC not defined	(+)	>0.70	(+)	the result in line with 3 Hypo's (3+)	(?)	No hypothesis defined	2(+), 2(?)
IPAS	USA	678	(?)	MIC not defined	(?)	Not all info	(+)	the result in line with 4 Hypo's (4+), Result not in line with 1 hypo's (1-)	(+)	The result is in accordance with the hypothesis	2(+), 2(?)
JeffSATIC	USA	1976	(?)	MIC not defined	(?)	Not all info	(+)	the result in line with 2 Hypo's (2+)	(?)	No hypothesis defined	1(+), 3(?)
ICCAS	Canada	584	(?)	MIC not defined	(?)	Not all info	(+)	the result in line with 2 Hypo's (2+)	(?)	No hypothesis defined	1(+), 3(?)
	USA	785	(?)	MIC not defined	(?)	Not all info	(+)	the result inline with 2 Hypo's (2+)	(?)	No hypothesis defined	1(+), 3(?)

Note: n = number of participant, MIC = minimal important change, hypo = hypothesis
Quality Criteria for Measurement Properties Ratings: (+) A positive rating indicates strong properties according to quality criteria using design and method; (?) An intermediate rating indicates some but not all aspects of psychometric are positive, or doubtful design or method: (-) A negative rating indicates that psychometric properties do not meet criteria despite adequate design and method.

Anna, A., et al. (2024)

Table 4a. Quality of the Evidence for Measurement Properties of the Instruments

Scale	SATP2C		SPICE-R		ATHCTS		IPAS		JeffSATIC	
	Rating	Qual.evid	Rating	Qual.evid	Rating	Qual.evid	Rating	Qual.evid	Rating	Qual.evid
Content validity:	?	Low	?	moderate	?	moderate	+	moderate	+	moderate
Relevance	?	Low	?	Low	?	moderate	+	high	+	moderate
Comprehensiveness	?	Low	+	high	+	high	+	high	+	moderate
Comprehensibility	-	very low	?	Low	?	low	?	low	+	moderate
Structural validity	?	Low	+	high	+	high	+	high	+	high
Internal consistency	+	moderate	+	high	+	high	+	high	+	high
Cross-cultural validity	+	moderate	?	low	+	high	+	high	?	moderate
Reliability	?	Low	+	high	+	high	?	low	?	moderate
Measurement error	?	Low	?	low	?	moderate	?	low	?	Low
Criterion validity	+	moderate	?	low	?	moderate	?	low	?	Low
Construct validity	-	very low	+	high	+	high	+	high	+	moderate
Responsiveness	?	low	?	low	+	moderate	+	high	?	Low

Note: COSMIN Rating definitions: "+" = sufficient; "?" "?" =indeterminate, "–", "–" = insufficient

members who work at Veteran Affairs Medical Centers (Heinemann et al., 1999). Subsequently, Curran et al. (2007) adapted the instrument to assess attitudes toward interprofessional healthcare teams among 1179 undergraduate health science students in Canada. The authors developed 14 items, comprising two subscales: quality of care and time constraints. Participants responded on a five-point Likert scale, ranging from one (strongly disagree) to five (strongly agree). In addition, Hayashi et al. (2012) used the ATHCTS to measure attitudes toward teamwork among 285 undergraduate students in Japan. They were still using 14 items; however, they divided it into three subscales: quality of care delivery, team efficiency, and patient-centered care.

Moreover, in 2014, Kim and Ko used the adapted ATHCTS with 14 items and two subscales, quality of care and time restriction, to measure attitudes toward teamwork among 288 graduate professional students in the USA. They scored on a five-point Likert scale (1= = strongly disagree, 5=strongly agree). A higher score on the scale indicates a more positive attitude towards interprofessional healthcare teams.

Furthermore, Norris et al. (2015) combined 16 items of the extended RIPLS (the Readiness Interprofessional Learning Scale) with 16 new items from the Interprofessional Education Competency (IPEC) Report. The questionnaire was developed with the objective of assessing interprofessional attitudes among a total of 1,549 undergraduate and graduate students enrolled in healthcare professions at the University of Utah Health Sciences Center. After the result of the EFA (Exploratory Factor Analysis), they named their tool the Interprofessional Attitudes Scale (IPAS). The instrument comprises 27 items, which are grouped into five subscales: teamwork, roles, and responsibility (TRR), patient-centredness (PC), interprofessional biases (IB), diversity and ethics (DE), and community-centredness (CC). Each item is rated on a five-point Likert scale, with 1 indicating strong disagreement and 5 indicating strong agreement.

Measure of the Interprofessional Education and Practice Perceptions
Originally, Fike et al. (2013) developed the Student Perceptions of Interprofessional Clinical Education (SPICE) instrument to assess perceptions of medical and pharmacy students regarding

Table 4b. Quality of the Evidence for Measurement Properties of the Instruments

Scale	RIPLS, HK		RIPLS,P		RIPLS,I		ICCAS		ICCAS-R	
	Rating	Qual.evid	Rating	Qual.evid	Rating	Qual.evid	Rating	Qual.evid	Rating	Qual.evid
Content validity	?	low	?	low	+	moderate	?	low	?	low
Relevance	?	low	?	low	+	high	?	low	?	low
Comprehensiveness	?	low	?	low	+	high	?	low	?	low
Comprehensibility	?	low	?	low	?	low	?	low	?	low
Structural validity	+	high	+	high	+	high	?	Low	?	Low
Internal consistency	+	high	+	high	+	high	+	high	+	high
Cross-cultural validity	?	low	?	low	+	high	?	Low	?	Low
Reliability	?	low	?	low	+	high	?	Low	?	Low
Measurement error	?	low	?	low	?	low	?	Low	?	Low
Criterion validity	?	low	?	low	+	high	?	Low	?	Low
Construct validity	+	high	+	high	+	high	+	high	+	high
Responsiveness	?	low	?	low	?	low	?	Low	?	Low

Note: COSMIN Rating definitions: "+" = sufficient, "?" =indeterminate, "–" = insufficient

interprofessional education (IPE). The scale comprises ten Likert-type items, with response options ranging from strongly disagree (1) to strongly agree (5). It is comprised of three subscales: interprofessional teamwork and team-based practice, roles/responsibilities for collaboration practice, and patient outcomes from collaboration practice.

The SPICE-Revised is the designation for the modification of SPICE. The 10-item SPICE-R scale was employed to assess perceptions of IPE and IPCP (interprofessional clinical practice) among 277 first-year students enrolled in a range of health professional degree programs, including the Bachelor of Science in Nursing, Master of Health Administration, Doctor of Optometry, and Doctor of Physical Therapy. Participants responded on a five-point Likert scale (1 = strongly disagree, 5 = strongly agree).

Measure of the Readiness for Interprofessional Learning

The Readiness for Interprofessional Learning Scale (RIPLS) was developed in 1998 by Parsell and Bligh. They validated the measurement to 914 respondents. The scale consisted of 19 items and the authors grouped them into two subscales: teamwork and collaboration (TWC) and professional identity (PI). Then, in 1999, Parsell and Bligh conducted the second study and yielded three subscales, including the TWC subscale (items 1-9), PI (items 10-16), and roles and responsibilities (RR) as the third subscale (item 17-19).

The RIPLS has been validated and adapted into different languages, including Swedish (2008), Japanese (2012), Indonesian (2014), French (2015), Danish (2016), and Persian (2019). They used the RIPLS self-report questionnaire to estimate the degree of students' readiness to engage in interprofessional learning activities before they graduate and work at any health service.

In Indonesia, Tyastuti et al. (2014) translated and cross-culturally adapted the original version of the RIPLS into the Indonesian language. They tested its reliability and validity on 755 first to third-year students from four health departments of the Faculty of Medicine and Health Sciences. The students answered 18 items on a five-point Likert scale (1 = strongly disagree, 5=strongly agree). Higher scores indicate more readiness for interprofessional learning. The authors omitted item 17 (the function

of nurses and therapists is mainly to support doctors) due to the absence of a therapist student participant in their study and the polysemy of the term "therapist" in Indonesian society.

Nevertheless, Ganotice and Chan (2018) conducted construct validation of the English version of RIPLS to 469 Chinese undergraduate students from two Hong Kong universities. They validated the 19-item English version with four subscales: teamwork and collaboration, negative professional identity, positive professional identity, and roles and responsibilities. The students answered all items on a five-point Likert scale, with higher scores reflecting a more significant endorsement of the readiness for interprofessional learning.

Further, in Iran, Ataollahi et al. (2019) translated and assessed the validity and reliability of the Persian version of the RIPLS. They measured 200 final-year medical students in Iran and used 19 items rated on a five-point Likert scale with the same subscales used by Ganotice and Chan (2018).

Measures of the International Collaborative Competency
The Interprofessional Collaboration Competency Attainment Survey (ICCAS) is used to measure the self-reported competencies of the IPE care program. The ICCAS was developed based on IPC competencies. These competencies are communication, collaboration, roles and responsibilities, a collaborative patient/family-centered approach, conflict management/resolution, and team functioning. So, using ICCAS, we can ask the learners to self-reflect on the changes in the level of competencies through an IPE intervention.

Archibald et al. (2014) conducted a study to assess the validity and reliability of the ICCAS. They developed a list of 20 items that corresponded to the six competencies, each item answered on a seven-point Likert scale (strongly disagree=1 to strongly agree=7), and an option to rate an item as 'not applicable. Then, 584 respondents from 15 interprofessional education programs in Canada and New Zealand participated in answering those 20 items.

In 2017, Schmitz et al. (2017) replicated a validation of the ICCAS. They wanted to find the extent of the ICCAS appropriate for their population and curriculum. They examined its validity to 783 students who enrolled in the Fundamentals in Interprofessional Communication and Collaboration (FIPPC) course. Before they distributed the ICCAS to the students, they made two changes to the instrument. Firstly, they changed the rating scale from a seven-point Likert type 'agree-disagree' format to a five-point, unbalanced, qualitative scale: poor (1), fair (2), good (3), very good (4), and excellent (5). Secondly, they added an item to assess the changes in the student's overall abilities during the FIPCC course. They used this result to assess the concurrent validity of ICCAS items. Before the administration, they conducted cognitive

interviews with several students to ensure that they understood their changes.

Discussion

This review has two aims: firstly, to systematically review and identify instruments available to measure IPE and IPC in health professional students, and secondly, to critically evaluate available instruments and provide recommendations about the most appropriate IPE/IPC instruments. With these aims, we analyzed the measurement properties and the use of various questionnaires assessing collaboration in interprofessional education and practice.

In almost all instruments developed to measure IPE, the attitudes, and perceptions measured are related to teamwork, roles, responsibilities and accountability, quality of care, and time constraints (Dominguez et al., 2015; Hojat et al., 2012; Kim & Ko, 2014). In contrast, IPAS measures teamwork, responsibilities, and relationships with patients and the community, including diversity and ethics (Norris et al., 2015).

In this case, IPAS measures diversity and ethical factors that play an important role in the process of collaborating with other professionals. Compared to other instruments, based on the assessment using COSMIN, the IPAS instrument has excellent structural validity, internal consistency, cross-cultural validity, reliability, collateral errors, and criterion validity. Both of them asked patients and experts for content validity and adequate relevance. This showed that the IPAS instrument has a very small risk of bias. These results are consistent with an international Delphi study conducted to develop the Consensus-Based Standards for the Selection of Health Measurement Instruments (COSMIN) initiative aimed at facilitating the selection of high-quality patient-reported outcome measures (PROMs) for research and clinical practice. High-quality systematic reviews can provide a comprehensive picture of the measurement properties of PROMs and support evidence-based recommendations in the selection of the most appropriate PROMs for a given purpose (Mokkink et al., 2010).

The instrument to measure student readiness (RIPLS) has been translated into Chinese (Ganotice & Chan, 2018), Persian (Ataollahi et al., 2019), and Indonesian (Tyastuti et al., 2014). From the evaluation results using COSMIN, the evaluation results related to content validation by experts on the Chinese and Persian readiness instruments were doubtful. However, this does not mean that the Chinese language version of the instrument was not relevant in terms of content. When reviewing using COSMIN, the reviewer only reviews according to the report written in the article being assessed.

Five items are assessed related to the relevance of the questions asked of patients and experts, including their analysis. For example, the first item asks: Was an appropriate method used to

ask patients whether each item is relevant to their experience with the condition? Then, the next item question asks: Was each item tested in an appropriate number of patients? For qualitative studies? For quantitative (survey) studies? And so on. The reviewer will provide a checklist of whether it is very good, adequate, doubtful, inadequate, or NA (Mokkink et al., 2010). Therefore, the results highly depend on how researchers report their research methods in the articles they write.

Methodological quality assessments indicated that some instruments were less rigorously developed and validated than those recommended by COSMIN and the Criteria for Quality of Measurement Properties (Mokkink et al., 2010; Terwee et al., 2007). Incomplete reporting of the item selection process and description of sample characteristics may limit content validity (Terwee et al., 2007). Some authors did not conduct confirmatory factor analysis (CFA) to explore the dimensionality of the questionnaire. Authors often reported the results of CFA and reliability analyses from previously published studies rather than analyzing both in their sample. This practice may need to be revised as a measurement model, given that factor loadings and reliability depend on sample data, measures, and missing items (Mokkink et al., 2010). Sometimes, authors analyze the reliability of IPE/IPC scores using correlations, so they cannot consider or assess systematic errors and concordance. It is essential to exercise caution when evaluating construct validity and sensitivity without testing specific hypotheses, as this can potentially lead to misleading conclusions. It is not uncommon for authors to propose alternative explanations for low correlations or small differences in means rather than concluding that the questionnaire may not be valid for its intended purpose (Mokkink et al., 2010; Terwee et al., 2007).

So, of the ten instruments reviewed using COSMIN, the Interprofessional Attitudes Scale (IPAS) was the most appropriate instrument to measure interprofessional competence compared to the other instruments. The IPAS can be used by health sciences educational institutions and other institutions to train people to work together in interprofessional teams.

Furthermore, the findings of this psychometric review offer a methodology for the selection and development of the most appropriate instrument for the assessment of IPE/IPC, based on the instrument's intended purpose, psychometric properties, and utility. Firstly, the authors should consider the use of a measurement model, which is reflective and formative, when developing a new instrument. A measurement model was defined as the construct of interest. Second, authors should adequately validate questionnaires for assessing IPE/IPC and correctly establish their reliability to ensure methodological quality. In this sense, translation and cross-cultural adaptation of questionnaires should be correctly validated. This effort is an attempt to provide the invariance of the construct.

Lastly, in instances where authors translated an instrument for measuring attitudes, perceptions, levels of readiness, or competencies related to IPE/IPC, they did not perform the requisite multi-group factor analysis to test construct invariance. Consequently, there is an elevated risk of biased results when comparing scores (Mokkink et al., 2010).

Strength and limitation
The COSMIN guidelines offer a number of advantages for the systematic review of patient-reported outcome measures (PROMs). They facilitate the selection of high-quality PROMs for research and clinical practice by providing a methodology for combining the methodological quality of studies on the nature of the measurement with the quality of the PROMs.

A systematic review of patient-reported outcome measures (PROMs) comprises ten sequential steps. The initial four steps pertain to the preparation and implementation of the literature search and the selection of pertinent studies. The subsequent six steps focus on evaluating the quality of eligible studies, measurement properties, and interpretability and feasibility aspects. Finally, steps nine and ten entail the formulation of recommendations and the reporting of the systematic review. Given the considerable number of items assessed, there is a risk of bias in the conclusions drawn since not all reports on PROM development adhere to the COSMIN guidelines. It is therefore recommended that researchers engaged in the development or translation of PROMs should make use of the COSMIN checklist to ensure the production of high-quality PROMs.

Conclusions

In general, from the review results, the authors found ten instruments (PROMs) that can be used to measure IPE and IPC in health professional students, namely the scale of attitudes toward physician–pharmacist collaboration (SATP2C), Student Perceptions of Physician-Pharmacist Interprofessional Clinical Education(SPICE), Attitudes toward interprofessional health care teams scale (ATHCTS), Readiness for Interprofessional Learning Scale (RIPLS) Chinese, Persian, and Indonesian versions, the Interprofessional Attitudes Scale (IPAS), the Jefferson Scale of Attitudes Toward Interprofessional Collaboration (JeffSATIC), and the interprofessional collaborative competency attainment survey (ICCAS).

Based on the evaluation results using the COSMIN guide, the Interprofessional Attitudes Scale (IPAS) is a PROM with adequate content validation and structural validity, internal consistency, cross-cultural validity, reliability, measurement error, and criterion validity, which are very good. It is

recommended that IPAS be used to measure the interprofessional attitudes of health professional students.

Declaration of Interest

Acknowledgment

Funding

Data Availability

The dataset in this study was accessible from the corresponding author upon reasonable request.

References

Archibald, D., Trumpower, D., & MacDonald, C. J. (2014). Validation of the interprofessional collaborative competency attainment survey (ICCAS). *Journal of Interprofessional Care, 28*(6), 553–558. https://doi.org/10.3109/13561820.2014.917407

Ataollahi, M., Amini, M., Delavari, S., Bazrafkan, L., & Jafari, P. (2019). Reliability and validity of the Persian version of readiness for inter-professional learning scale. *International Journal of Medical Education, 10*, 203–207. https://doi.org/10.5116/ijme.5da4.37c2

Curran, V. R., Sharpe, D., & Forristall, J. (2007). Attitudes of health sciences faculty members towards interprofessional teamwork and education. *Medical Education, 41*(9), 892–896. https://doi.org/10.1111/j.1365-2923.2007.02823.x

Dominguez, D. G., Fike, D. S., Maclaughlin, E. J., & Zorek, J. A. (2015). A comparison of the validity of two instruments assessing health professional student perceptions of interprofessional education and practice. *Journal of Interprofessional Care, 29*(2), 144–149. https://doi.org/10.3109/13561820.2014.947360

Fike, D. S., Zorek, J. A., MacLaughlin, A. A., Samiuddin, M., Young, R. B., & MacLaughlin, E. J. (2013). Development and validation of the student perceptions of physician-pharmacist interprofessional clinical education (SPICE) instrument. *American Journal of Pharmaceutical Education, 77*(9), 1–8. https://doi.org/10.5688/ajpe779190

Ganotice, F. A., & Chan, L. K. (2018). Construct validation of the English version of Readiness for Interprofessional Learning Scale (RIPLS): Are Chinese undergraduate students ready for 'shared learning'? *Journal of Interprofessional Care, 32*(1), 69–74. https://doi.org/10.1080/13561820.2017.1359508

Hayashi, T., Shinozaki, H., Makino, T., Ogawara, H., Asakawa, Y., Iwasaki, K., Matsuda, T., Abe, Y., Tozato, F., Koizumi, M., Yasukawa, T., Lee, B., Hayashi, K., & Watanabe, H. (2012). Changes in attitudes toward interprofessional health care teams and education in the first- and third-year undergraduate students. *Journal of Interprofessional Care, 26*(2), 100–107. https://doi.org/10.3109/13561820.2011.644355

Heinemann, G. D., Schmitt, M. H., Farrell, M. P., & Brallier, S.A. (1999). Development of an attitudes toward health care teams scale. *Evaluation and the Health Professions, 22*(1), 123–142. https://doi.org/10.1177/01632789922034202

Hojat, M., & Herman, M. W. (1985). Developing an instrument to measure attitudes toward nurses: Preliminary psychometric findings. *Psychological Reports, 56*(2), 571–579. https://doi.org/10.2466/pr0.1985.56.2.571

Hojat, M., Spandorfer, J., Isenberg, G. A., Vergare, M. J., Fassihi, R., & Gonnella, J. S. (2012). Psychometrics of the scale of attitudes toward physician-pharmacist collaboration: A study with medical students. *Medical Teacher, 34*(12), 833–837. https://doi.org/10.3109/0142159X.2012.714877

Hojat, M., Ward, J., Spandorfer, J., Arenson, C., Van Winkle, L. J., & Williams, B. (2014). The Jefferson Scale of Attitudes Toward Interprofessional Collaboration (JeffSATIC): Development and multi-institution psychometric data. *Journal of Interprofessional Care, 29*(3), 238–244. https://doi.org/10.3109/13561820.2014.962129

Kajander-Unkuri, S., Meretoja, R., Katajisto, J., Saarikoski, M., Salminen, L., Suhonen, R., & Leino-Kilpi, H. (2014). Self-assessed level of competence of graduating nursing students and factors related to it. *Nurse Education Today, 34*(5), 795–801. https://doi.org/10.1016/j.nedt.2013.08.009

Kim, K., & Ko, J. (2014). Attitudes toward interprofessional health care teams scale: A confirmatory factor analysis. *Journal of Interprofessional Care, 28*(2), 149–154. https://doi.org/10.3109/13561820.2013.857645

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ (Online), 339*(7716), 332–336. https://doi.org/10.1136/bmj.b2535

Mokkink, L. B., de Vet, H. C. W., Prinsen, C. A., & Terwee, C. B. (2018). COSMIN Methodology for Conducting Systematic Reviews of Patient-

Reported Outcome Measures (PROMs). *Encyclopedia of Quality of Life and Well-Being Research, February,* 1–3. https://doi.org/10.1007/978-3-319-69909-7_2972-2

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & De Vet, H. C. W. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research, 19*(4), 539–549. https://doi.org/10.1007/s11136-010-9606-8

Norris, J., Carpenter, J. G., Eaton, J., Guo, J.-W., Lassche, M., Pett, M. A., & Blumenthal, D. K. (2015). The development and validation of the interprofessional attitudes scale. *Academic Medicine, 90*(10), 1394–1400. https://doi.org/10.1097/acm.0000000000000764

Oates, M., & Davidson, M. (2015). A critical appraisal of instruments to measure outcomes of interprofessional education. *Medical Education, 49*(4), 386–398. https://doi.org/10.1111/medu.12681

Schmitz, C. C., Radosevich, D. M., Jardine, P., MacDonald, C. J., Trumpower, D., & Archibald, D. (2017). The Interprofessional Collaborative Competency Attainment Survey (ICCAS): A replication validation study. *Journal of Interprofessional Care, 31*(1), 28–34. https://doi.org/10.1080/13561820.2016.1233096

Terwee, C. B., Bot, S. D. M., de Boer, M. R., van der Windt, D. A. W. M., Knol, D. L., Dekker, J., Bouter, L. M., & de Vet, H. C. W. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology, 60*(1), 34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012

Terwee, C. B., Mokkink, L. B., Knol, D. L., Ostelo, R. W. J. G., Bouter, L. M., & De Vet, H. C. W. (2012). Rating the methodological quality in systematic reviews of studies on measurement properties: A scoring system for the COSMIN checklist. *Quality of Life Research, 21*(4), 651–657. https://doi.org/10.1007/s11136-011-9960-1

Tibi, M. H. (2015). Improving collaborative skills by computer science students through structured discussion forums. *Journal of Technologies in Education, 10*(3–4), 27–41. https://doi.org/10.18848/2381-9243/cgp/v10i3-4/56460

Tyastuti, D., Onishi, H., Ekayanti, F., & Kitamura, K. (2014). Psychometric item analysis and validation of the Indonesian version of the Readiness for Interprofessional Learning Scale (RIPLS). *Journal of Interprofessional Care, 28*(5), 426–432. https://doi.org/10.3109/13561820.2014.907778

World Health Organization [WHO]. (2010). Framework for action on interprofessional education & collaborative practice. https://www.who.int/publications/i/item/framework-for-action-on-interprofessional-education-collaborative-practice