Revista Latinoamericana de Psicología

https://revistalatinoamericanadepsicologia.konradlorenz.edu.co/

REVIEW

Assessing face-to-face dating violence in Ibero-America: Systematic review and meta-analysis of measurement instruments

Alhena L Alfaro-Urquiola [a,b]*, Luis Burgos-Benavides [b], Andrés Ramírez [c], Francisco Javier Herrero [b], Francisco Javier Rodríguez-Díaz [b]

[a] *Universidad Católica Boliviana "San Pablo", Bolivia*
[b] *Universidad de Oviedo, España*
[c] *Universidad Politécnica Salesiana, Cuenca, Ecuador*

Received 21 May 2024; ***accepted*** 16 March 2025

Abstract | **Introduction:** Dating violence is a significant social concern, requiring reliable measurement tools for accurate assessment. This study conducts a systematic review and meta-analysis to evaluate and synthesise existing instruments used to measure face-to-face dating violence employed in Ibero-America. **Method:** The study followed PRISMA guidelines, conducting a comprehensive search from December 2022 to April 2023 across SCOPUS, PubMed, PsychINFO, and Web of Science databases. Inclusion criteria covered adolescents and young adults, singles, and instruments validating dating violence measurements in Ibero-America, excluding precursors and related factors. **Results:** The search yielded 247 articles, after depuration, 21 studies were included, revealing 16 dating violence assessment instruments. Spain was the primary source; predominantly non-clinical samples were used (81% students). Many scales demonstrated adequate reliability (α > 0.7) and exhibited strong construct validity supported by confirmatory factor analyses. The significant variability across studies hinders comparability. Although no scale meets all the evaluated parameters, the DVQ-VP and CARPA stand out in terms of validity and reliability. **Conclusions:** While many scales are valid and suitable in terms of validity, the need for future research employing current criteria to assess these aspects is emphasised. Additionally, there is a call for further exploration of predictive and concurrent validity, as well as gender invariance.

Keywords: Meta-analysis, dating violence, reliability, adolescents, psychometrics

Evaluando la violencia presencial en el noviazgo en Iberoamérica: revisión sistemática y metaanálisis de instrumentos de medida

Resumen **Introducción:** La violencia en el noviazgo es un tema de gran interés social y por tanto su evaluación requiere herramientas de medición confiables para una mayor precisión. Este estudio realiza una revisión sistemática y un metaanálisis de los instrumentos existentes utilizados para medir la violencia presencial en el noviazgo en Iberoamérica. **Método:** Se siguieron las directrices PRISMA, llevando a cabo una búsqueda exhaustiva entre diciembre de 2022 y abril de 2023 en las bases de datos SCOPUS, PubMed, PsycINFO y Web of Science. Los criterios de inclusión abarcaron estudios con adolescentes y adultos jóvenes, personas solteras y artículos que validaran instrumentos de medida para la violencia en el noviazgo en Iberoamérica, excluyendo factores precursores y relacionados. **Resultados:** La búsqueda inicial identificó

* Corresponding author.
E-mail: agathalhe@gmail.com

https://doi.org/10.14349/rlp.2024.v56.28

247 artículos; tras el proceso de depuración, se incluyeron 21 estudios, que revelaron la existencia de 16 instrumentos de evaluación de la violencia en el noviazgo. España fue el país con mayor número de estudios, y en su mayoría se emplearon muestras no clínicas (81 % estudiantes). Muchas escalas demostraron una fiabilidad adecuada (α > 0.7) y presentaron una sólida validez de constructo respaldada por análisis factorial confirmatorio. Sin embargo, la variabilidad significativa entre los estudios dificulta la comparación. Aunque ninguna escala cumple con todos los parámetros evaluados, el DVQ-VP y el CARPA se destacan en términos de validez y fiabilidad. **Conclusiones:** Si bien muchas escalas son válidas y adecuadas en términos de validez, se enfatiza la necesidad de futuras investigaciones que utilicen criterios actuales para evaluar estos aspectos. Además, se destaca la importancia de explorar con mayor profundidad la validez predictiva y concurrente, así como la invarianza de género.

Palabras clave: Metaanálisis, violencia en el noviazgo, fiabilidad, adolescentes, psicometría

Terms such as intimate partner violence, dating violence or couple violence can be read more and more in newspapers and articles, either to report its prevalence, analyse risk or protective factors or intervene with victims and aggressors (Juarros-Basterretxea et al., 2022; López-Barranco et al., 2022b; Torres et al., 2022). The reason: there is a greater number of people who experience or have experienced some type of violence within their relationship, especially women (Tarriño-Concejero at al., 2023). Systematic reviews of its prevalence have found very different figures, many of them depending on the context in which it was evaluated (culture, age, sex, role, sample, etc.), the definition used (violence in courtship, intimate partner violence, teen dating violence, gender-based violence, cyberviolence, cyberviolence in courtship, etc.) and also the scale used (Exner-Cortens et al., 2016; Jennings et al., 2017).

This study focuses on dating violence, understood as the set of attitudes, behaviours and forms of relationship in which violence, threats or provocations occur with the purpose of generating control or emotional, verbal, psychological, physical and/or sexual damage in the context of a dating relationship in the absence of cohabitation, children or economic dependence (Gracia-Leiva et al., 2019; Marcos et al., 2023). The members of the couple may be adolescents or young adults (Jennings et al., 2017). As time has passed and technology and access to it have advanced, more recent definitions mention that it can occur in person, online or through technology (Centres for Disease Control and Prevention, 2016, 2023; Fernet et al., 2019).

A proper evaluation of the phenomenon hinges on the scale's quality and its psychometric properties. Yanez-Peñunuri et al. (2019) underscored this in their review on dating violence questionnaires in Ibero-America. They identified 22 articles in other systematic reviews utilising the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) and Terwee et al. (2007) criteria for cross-cultural validation. Although these instruments demonstrated internal consistency, content, and construct validity, they lacked criterion validity, reproducibility, and sensitivity information. In the systematic review on the psychometric properties of instruments that qualitatively assess dating violence, prepared by Tarriño-Concejero et al. (2023), 29 scales were found, of which only three met the criteria proposed by the COSMIN guidelines.

Although the aforementioned reviews and recent studies (Alexander et al., 2022; Martínez-Soto & Ibabe, 2022; Tarriño-Concejero et al., 2023) employed COSMIN standards from the COSMIN Risk of Bias checklist (Mokkink et al., 2018), these guidelines were primarily designed for Patient-Reported Outcome Measures (PROMs) which have particular characteristics not seen in other settings (Prinsen et al., 2018; Terwee et al., 2012) and which, as is the case with violence, may differ depending on whether the sample is community, clinical, or prison-based (Love et al., 2020). However, many of the studies included in these studies were conducted with student populations, which could have introduced biases when evaluating these publications which could affect the generalisability of the results. Advances in science and technology in data analysis have facilitated greater precision in the treatment of multivariate phenomena and non-normal distributions. Violence, in particular, is a phenomenon that has these characteristics, since it is inherently complex and does not follow normal distribution patterns. As a result, many of the validity and reliability criteria previously considered appropriate for evaluating instruments have become obsolete or are not recommended for current measurement (Ferrando et al., 2022; Raykov, 1997). Additionally, although the previous works mentioned make an exhaustive review of the existing literature, they lack a meta-analysis that allows the quantitative comparison of the studies (Sánchez-Meca et al., 2013).

Based on the above, the objective of the systematic review is to evaluate the methodological quality of the studies on the measurement properties including the validity, reliability and internal consistency of face-to-face dating violence instruments for the evaluation of Ibero-American adolescents and emerging adults using current criteria and meta-analysis.

Method

Design

A systematic review was conducted to assess the psychometric properties of dating violence measurement instruments following the recommendations in Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA, Page et al., 2021).

Search strategy

The literature search was conducted between December 2022 and April 2023 by two independent researchers, consulting SCOPUS, PubMed, PsychINFO and Web of Science databases. All original studies, regardless of the date of publication, were considered in the search using the eligibility criteria.

The inclusion criteria were: (a) *Participants*: adolescents and young adults between 15 and 29 years of age. If the participants are younger or older but the age range is within the afore mentioned criteria, the studies will also be included in the review. Due to the particularities of dating violence, participants must be single and not live with their partner. (b) *Studies*: studies that describe the design or validation of instruments to measure any dimension of direct (face-to-face) dating violence but may or may not include items of cyber violence. (c) *Instrument*: self-reported quantitative instrument. Were excluded all the studies than had instruments that measure precursors of dating violence such as beliefs, perceptions or related factors, whether they were protective or risk factors.

Using that criteria connected with Boolean operators (AND, OR, AND NOT) and Article as a document type, the following common search strategy was used for all databases for Title, abstract and keywords: "dating violence" OR "gender-based violence" OR "intimate partner violence" OR "violence against women" OR "relationship violence" OR "dating abuse" OR "dating partner aggression" (Topic) and adolescent** OR teenager OR "young adult*" OR youth (Topic) and "validation study" OR "validation scale" OR "instrument validation" OR validation OR "psychometric properties" OR "validity assessment" OR "inventory validation" (Topic) and Latin* OR Latin-American** OR Caribbean OR Portugal OR Portuguese OR Hispanic OR Brazilian OR Spain OR Argentina OR Bolivia OR Chile OR Colombia OR "Costa Rica" OR Cuba OR equator OR "El Salvador" OR Guatemalan OR Honduras OR Mexico OR Nicaragua OR Panamá´ OR Paraguay OR Perú OR "Puerto Rico" OR "República Dominicana" OR Uruguay OR Venezuela (Topic).

Since the search returned many results (1,230) that included violence within marriage or violence towards children, the search was limited by making the following exclusion NOT children OR child OR toddler OR infant OR "married couple" OR baby OR marriage OR wife OR husband or married (figure 1). All articles whose scale exclusively evaluated cyber violence were discarded manually.

During the article review stage, the inclusion of a new article was suggested, which, although it did not meet the inclusion criteria based on its keywords, did meet them in terms of its methodology.

Meta-analysis strategy

A meta-analysis was conducted to examine the reliability of the dating violence scale, synthesising results to infer test reliability across samples. Effect sizes and alpha variances were calculated using the Hakstian-Whalen transformation to normalise reliability coefficients (Hakstian & Whalen, 1976). Heterogeneity was assessed using Cochran's Q, I^2, H^2, and τ^2 (Hedges & Olkin, 1985; Higgins & Thompson, 2002). For experienced violence, the Average Variance Extracted (AVE) served as a moderator.

Results

The search strategy resulted in 247 articles, from this, 114 were duplicated, as shown in Figure 1. After the analysis of the inclusion and exclusion criteria, 20 articles were included in the present study; 15 instruments were identified. After the article review process, a new article and scale were added (Figure 1).

Description of the instruments and studies

Of the studies, 57.14% were conducted in Spain, 19.05% in Mexico, and 4.76% each in Colombia, Chile, Portugal, the U.S., and Italy. All used non-clinical samples, with 81% focusing on students. Only three employed probabilistic sampling. Most studies (57%) had more women, 10% included only women, 33% had balanced genders, and 10% had more men. Two studies reported non-binary gender or other sexual orientations (Presaghi, et al., 2015; Soriano-Ayala et al., 2021). Sample size ranged from 398 (Javier-Juárez et al., 2022) to 6,138 (Rodríguez-Díaz et al., 2017). The studies were published between 2007 and 2022; the year with the highest production was 2021 (n = 4) and the lowest were 2007, 2015 and 2016 (n = 1).

An important part of the content validity of a scale is related to the fact that its items are not only correct but also understandable. In this sense, double translations were made (English-Spanish-English; English-Portuguese-English) in the validation of the TDV-VP, PMWI-SF and DVQ-R (2016 version), simple translations in the validation of the DVQ (2021 and 2015) and more in-depth linguistic reviews for the validation of the M-CTS (2019). In the CMN(2014) and M-CTS (2007) evaluations, although the process followed is not specified, it is mentioned that an adaptation to the language was made (Table 1 and Table 2).

The content of the items was adapted to the cultural context by 11 of the 21 studies. Either with one or more of the processes detailed below: a) using the Guidelines for the translation and adaptation of Tests proposed by the International Test Commission (ITC, 2017); b) semantic adaptations (TDV-VP); c) working with groups of experts both on the subject and in the language or in research methods (MSDV 2.0, CARPA; 2021's DVQ, TDV-VP, EMVN, PMWI-SF and M-CTS); d) discussion or focus groups (DVQ, 2021); e) redrafting of items after a pilot test (CMN), based on other scales (ESVIGA) or f) specifically created for the context from other scales (TDV-VP and VADRI-MX).

The differential analysis of the variables indicates that the age of the participants is within a range from 11 (Calvete et al., 2021) to 55 (Começanha & Maia, 2018). Regarding to the nationality of the participants, 57% of the studies (n = 12) included Spaniards and 43% Latin American population, while 2 studies took into account Portuguese and another 2 studies related to Italians.The original tests were created between 1979 (CTS precursor

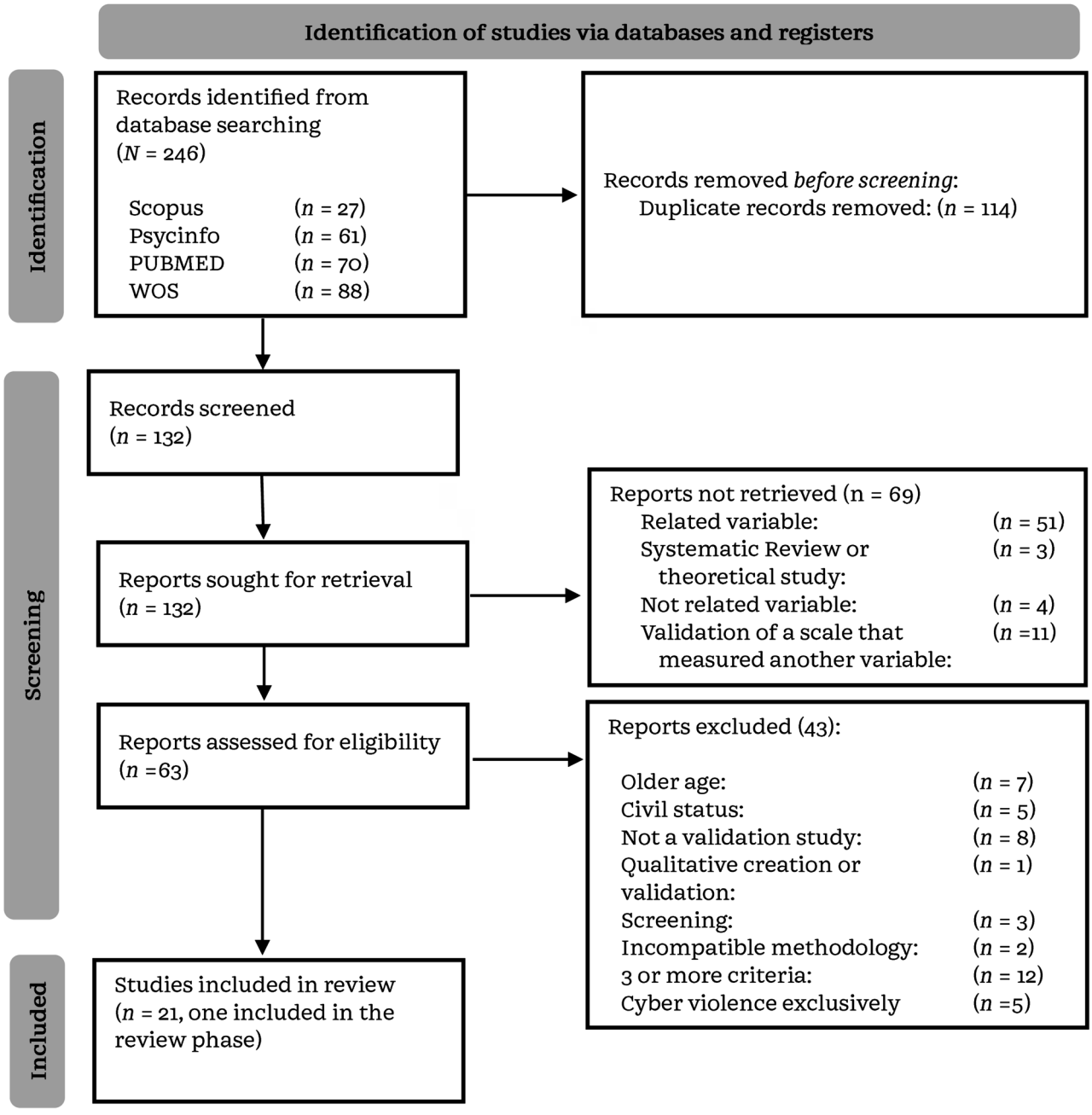

Figure 1. Flowchart of the Review Process according to The PRISMA 2020 statement: an updated guideline for reporting systematic review

Table 1. Authors, journal, year, country and scales used in the systematic review

N°	Author	Journal	Year	Country	Scale used
1	García-Carpintero-Muñoz et al.	Journal of Advanced Nursing	2022	Spain	Multidimensional Scale of Dating Violence -MSDV

(*Continued*)

N°	Author	Journal	Year	Country	Scale used
2	Lopez-Barranco et al.	International Journal of Environmental Research and Public Health	2022b	Spain	Conflict in Adolescent Dating Relationships Inventory -CADRI
3	Rodriguez-Franco et al.	International Journal of Clinical and Health Psychology	2022	Spain	Dating Violence Questionnaire for Victimisation and Perpetration - DVQ-VP
4	Calvete et al.	Revista de Psicología Clínica con Niños y Adolescentes	2021	Spain	Cuestionario de Abuso en Relaciones de Pareja de Adolescentes - CARPA
5	Lara and Lopez-Cepero	Journal of Interpersonal Violence	2021	Chile	Dating Violence Questionnaire - DVQ
6	Pacheco et al.	Revista Iberoamericana de Diagnóstico y Evaluación	2021	Colombia	Conflict in Adolescent Dating Relationships Inventory -CADRI
7	Soriano-Ayala et al.	International Journal of Environmental Research and Public Health	2021	Spain	Scale TDV-VP Teen Dating Violence. Victimisation and Perpetration
8	Aizpitarte and Rojas-Solís	International Journal of Psychological Research	2019	Mexico	Violence in Adolescents' Dating Relationships Inventory for Mexican Youth (VADRI-MX)
9	Ronzón-Tirado et al.	Frontiers in Psychology	2019	Mexico	Modified Conflict Tactics Scale – M-CTS
10	Garcia-Carpintero et al.	Gaceta Sanitaria	2018	Spain	Multidimensional Scale Dating Violence -MSDV
11	Começanha and Maia	Violence and Victims	2018	Portugal	Psychological Maltreatment of Women Inventory—Short Version
12	Penado-Abilleira and Rodicio-García	Anuario de Psicología Jurídica	2018	Spain	Adolescent Gender-Based Violence Scale
13	Muñoz-Rivas et al.	Behavioural Psychology/ Psicologia Conductual	2017	Spain	Sexual Coercion Scale -ECS
14	Rodríguez-Díaz et al.	International Journal of Clinical and Health Psychology	2017	Spain	Dating Violence Questionnaire-R – DVQ-R
15	López-Cepero et al.	Violence Vict	2016	United States	Dating Violence Questionnaire – DVQ
16	Presaghi et al.	Plos One	2015	Italy	Dating Violence Questionnaire-Italian version - DVQ (it)
17	Benitez Muñoz and Muñoz Bandera	Universitas Psychologica	2014	Spain	Conflict in Adolescent Dating Relationships Inventory – CADRI
18	Osorio-Guzmán	Salud Pública de México	2014	Mexico	Questionnaire Dating Abuse Italy-Mexico binational version- CMN
19	Viejo et al.	Anales de Psicologia	2014	Spain	Conflict Tactics Scale -CTS
20	Muñoz-Rivas et al.	Psicothema	2007	Spain	Modified version of the Conflict Tactics Scale -M-CTS
21	Javier-Juárez et al.	Cadernos de Saúde Pública	2022	Mexico	Violence in Adolescents' Dating Relationships Inventory - VADRI

Table 2. Characteristics of the samples and results of the articles used in the systematic review

N°	Year	N	Men	Woman	Sample	Age	Ethnicity	Sample type	Victimization	Perpetration
1	2022	1091	15%	85%	Non probabilistic by convenience	18-24 *M* = 20.1 *SD* = 1.67	96% Spain, 4% Italy, Marroquies, Brazil and Portugal	General, students	Sexual = 6.2% -Cyberbullying 24.7%	Sexual = 0.2% -Cyberbullying 10.6%
2	2022	976	82%	16.80%	Non probabilistic, intentional	19-25 *M* = 21.7 *SD* = 1.8	Spain	General, students	W: Physical = 11.2% Verbal-emotional abuse 83.7%; *M*: Physical = 15.2% Verbal emotional abuse = 82.9%	W: Physical-12%-Verbal-emotional abuse 89.9% M: Physical 4.9% verbal-emotional abuse 83.5%
3	2022	1232 (616 couples)	50%	50%	-	18-26 (*M* = 21.07, *SD* = 2.29)	Spain	General	PH: *M* = 0.86 *SD* = 1.19; S: M = 0.52 *SD* = 1.17; H: *M* = 1.35 *SD* = 0.97; De: *M* = 1.45 *SD* = 1.77; Coe = 1.24 *SD* = 1.64	PH: M = 0.72 *SD* = 1.05; S: M = 0.35 *SD* = 0.93; H: M = 0.74 *SD* = 1.02; De: M = 1.21 *SD* = 1.41; Coe = 1.05 *SD* = 1.40
4	2021	886		51.70%	Non probabilistic, incidental	11 a 18 *M* = 14.49 *SD* = 1.45	Spain	General, students	*t* = 10.5 (online)-34.5% (verbal) Total = 42.6% M = 35% W = 46.9%	*t* = 6.7 (online)-41.5% (verbal) Total = 52.5% M = 39.5% W = 64.6
5	2021	846	36.10%	63.90%	-	14-24 *M* = 17.87 *SD* = 2.72	Chile	General, students	Total:D *M* = 0.38 *SD* = 0.48 H *M* = 0.17 *SD* = 0.38 S *M* = 0.16 *SD* = 0.40 C *M* = 0.41 *SD* = 0.50 Ph *M* = 0.09 *SD* = 0.27 G *M* = 0.21 *SD* = 0.39 EP *M* = 0.25 *SD* = 0.45 I *M* = 0.09 *SD* = 0.29	
6	2021	2058	47.13%	52.86	-	12-19 *M* = 16.15 *SD* = 1.61	Colombia	General	Fis: *M* = 0.62 *SD* = 1.51; sex X)1.2 *DS* = 1.67; Rel. *M* = 0.79 *SD* = 1.46; verbal emotional *M* = 6.77 *SD* = 5.46; threats M = 0.72 *SD* = 1.28 Total: *M* = 10.09 *SD* = 8.90	

(*Continued*)

N°	Year	N	Men	Woman	Sample	Age	Ethnicity	Sample type	Victimization	Perpetration
7	2021	422	26.06%	73.70%	Non probabilistic	13-21 Female: *M* = 18.068 *DT* = 1.791 Male: *M* = 17.769 *DT* = 1.660	97,4% Spain; 2.6% were born in other countries: Romania, England, Germany, Poland, France, Mexico and Cuba	General		
8	2019	1055	51.90%	48.10%	Non probabilistic	14-22 *M* = 17.66 *SD* = 1.95	México	General, students	-	-
9	2019	1861	42.5%	57.50%	Non probabilistic	12 -18 *M* = 15.5 *SD* 1.39	México	Students	V *M* = 0.03-4.99	P M = 0.02-5.12;
10	2018	477	-	-	Probabilistic- random sampling stratified by areas of knowledge	-	Spain	General	-	-
11	2018	506	0	100%		18-55 *M* = 23.47 *SD* = 5.37	Portugal	General	-	-
12	2018	701	46.10%	53.60%		13-18 *M* = 16.14; *SD* = 2.25	Spain	Students	-	-
13	2017	3665	45.30%	54.70%	-	16-24 *M* = 19.92 *SD* = 2.47	Spain	General, students	-	-
14	2017	6138	39.6	60.40%	-	15-26 *M* = 18.5; *SD* = 2.09	Spain	General, students	-	-
15	2016	859	33.3	66.60%	-	18-26(*M* = 19 years; *SD* = 1.5 years).	55% White, 22% as African American, 12% as Asian, whereas 11% other identities. 9% Hispanic	Students	-	-
16	2015	418, final 139	22%	74%	-	16-26 *M* = 22 *SD* 1.88	Italy	Students	*M* = 0.1 (instrumental female. late adolescent) -*M* = 3.56 (coercion male. late adolescents)	-
17	2014	571	29.90%	70.10%	Probabilistic, simple random sampling	17-21 *M* = 18.76 *SD* = 1.204	Spain	Students	-	-
18	2014	2157	0	100%	Non probabilistic, incidental	14-33; *M* = 18.81 *SD* = 2.57	México	General, students	-	-
19	2014	2687	45.80%	54.20%	Stratified	15-21 *M* = 16.85 *SD* = 1.24	Spain	Students	F = 1-23.7% M = 1.9-12.5%	F = 0.9%-15.8% M = 2.7%-20.5%
20	2007	5355	63%	36.70%	-	16-26 *M* = 19.67 *SD* = 2.83	Spain	Students and workers		

(*Continued*)

N°	Year	N	Men	Woman	Sample	Age	Ethnicity	Sample type	Victimization	Perpetration
21	2022	398	37.2%	62.8%	Non probabilistic	15-18 *M* = 16.1 *SD* = 1	México	Students	*M*: F = 3.72; M = 6.02 *SD*: F = 6.6; M = 7.56	

From Men to Age Columns: *M* = mean; in victimization and perpetration columns: *M* = male; *F* = female

of M-CTS) and 2022 (MSDV 2.0 and DVQ-VP); three of the scales were used in more than one study: the CADRI and DVQ three times, and the M-CTS twice. As shown on Table 3, the scales ranged from 10 to 57 items distributed in two (PMWI-SF, ECS, CTS) to twelve factors (CADRI). Mostly evaluating the violence committed and the violence experienced as a victim (CADRI, DVQ-VP, MSDV 2.0, CARPA, TDV-VP, VADRI-MX, ECS, M-CTS, ESVIGA and CTS). Of all the instruments found, the MSDV 2.0, CARPA, EMVN and ESVIGA include at least one specific dimension linked to cyberviolence; the TDV-VP and VADRI-MX have items that assess the presence of some type of violence exercised through electronic devices and the DVQ-VP and PMWISV have one item related to it.

Results of measurement properties

For a better understanding of the results, the following abbreviated scale names include the validation year when multiple studies exist. All the scales except the CTS reported at least one reliability indicator, most (*n* = 13) reported Cronbach's alpha as an indicator, four used ordinal alpha, four McDonald's omega and one used Raykov's omega. At the level of the global scale or the scale divided according to the violence experienced or perpetrated, all the scales except the CADRI (only one of the three studies report a global score), ECS (both subscales) and the M-CTS (violence perpetrated) had values higher than .80 in at least one of the studies where they were analysed. It is important to mention that in the case of the TDV-VP, CTS and DVQ-VP, no overall reliability data was presented, and it was only analyzed based on its factors.

As can be seen on Table 3, regarding the reliability analysis by factors, the α values oscillated between .306 (M-CTS) and .96 (VADRI). The scales with the best scores were the ESVIGA, PMWI-SF and VADRI-MX; using ordinal alpha or omega, the best scores were, the DVQ, VADRI and CARPA.

Although most of the studies use Cronbach's alpha as an indicator of reliability of the scale, either globally or for the subscales, it should be noted that in all cases they are Likert-type scales (only the VADRI, VADRI- MX and EMVN have scales of six or more points), multidimensional and with data whose distribution is not normal, so it would not be the most appropriate indicator (Gadermann et al., 2012; Kelley & Pornprasertmanit, 2016; Raykov, 1997; Zumbo et al., 2007). In this sense, indicators such as McDonald's Omega, Raykov's Omega or even Ordinal alpha would be more appropriate and were only reported by the DVQ (ordinal alpha and Omega), CARPA (ordinal alpha and Omega), M-CTS (Omega), VADRI (Omega) and DVQ-VP (Omega). In the previously mentioned scales, all except the M-CTS presented data equal to or greater than .78, either for the global scale or for all its factors. Most of the studies (*n* = 17) used Cronbach's alpha as a measure of internal consistency, with only three studies additionally reporting the correlation between the factors or between the test items and the total (CARPA, EMVN, CMN), this last criterion was referenced by sex in only one study, the validation of the CTS (see Table 3).

To evaluate this criterion, as part of the present study, based on the standardised factor loadings provided by the articles, the composite reliability (CR) was calculated for both the global scales and for each of their factors. Obtaining general CR values ranging from .77 (ECS) to .98 (the CARPA whose validation study was the only one that presented composite reliability values); all the scales except the ECS had values greater than .70. In the analysis by dimension, there were more instruments that reported data below .7 in at least one of its dimensions (the CADRI, MSDV 2.0, Scale TDV-VP, ESVIGA, DVQ-R, M-CTS and CTS). If the recommendation of Fornell and Larcker (1981) is followed (> 0.6), at the global scale level or divided according to perpetration and victimisation, all the scales for which the calculation was made except the CTS obtained acceptable values and at the subscale level the DVQ-VP, DVQ-R, PMWI-SF, VADRI-MX and TDV-VP (see Table 3).

Only the MSDV 2.0 and PMVW-SF provided specificity and sensitivity data, but comparison is hindered by differing presentation methods: the former by factor, the latter by percentage for the global scale.

The reviewed studies presented data from exploratory factor analysis for 9 of the scales (CARPA, TDV-VP, VADRI-MX, EMVN, ESVIGA, ECS, DVQ-R, M-CTS, and CTS) and confirmatory factor analysis for all. All the studies found included the RMSEA, 19 the CFI, 14 the ratio between the chi square and degrees of freedom, 13 the chi square, 10 the TLI, 5 the GFI, 5 the AGFI, 3 the NFI, 2 the SRMR, 3 IFI, 2 R2 and other indicators (WRMR, PGFI, CMIN/gl, S-Bc²/gl, RFI, PNFI, MFI, NNFI, PCFI, AIC, ECVI) were analysed by two or fewer studies. Although many more model fit indicators were found than those presented on Table 4, they were not considered by most studies, so those that were used to a greater extent were extracted.

In the case of incremental adjustment indices, the CFI (≥. 90 or ≥ .95), GFI (≥ .93) and TLI (≥ .90,) were considered. For the absolute fit indices, RMSEA (≤ 0.05) and the SRMR (< 0.08) were taken into account (Cho et al., 2020; Lai, 2021; Xia & Yang, 2019). Although in the scales that were reviewed by more than one study, different results

Table 3. Scale, studies, factors, items and reliability of the studies that make up the systematic review

Scale	Original Author	N° studies	Items*	N° Factors	Violence	Alpha	Alpha dimensions	Other	CR	CR dimensions
CADRI	Fernández-Fuertes et al. (2006)	3	21-34	6 or 5x2	V&P; Off	V = .84 (1) P = .73	V = .474-.9 P = .401-.84	-	**V = .831 P = .783**	V = .267-.801 P = .502-.978
DVQ	Rodriguez-Franco et al. (2010)	3	42-46	8	V; Off	**.96**	.58-.893	$\alpha_{ordinal}$: **.82-.94** ω **= .83-.94**	**.983-.985**	.796-.949
MSD V 2.0	García-Carpintero et al. (2022)	1	2x18	2x5	V&P; Off & On	**V = .879 P = .802**	V = .703-.879 P = .702-.869	-	**V = .941 P = .917**	V = .368-.641 P = .244-.670
CARPA	Calvete et al. (2021)	1	2X24	2X5	V&P; Off&On	**V = .96 P = .95**	V = .79-.94 P = .74-.97	$\alpha_{ordinal}$: **V = .97 P = .96**	**V = .98 P = .97 ***	**V = .82-.94 P = .76-.96**
TDV-VP	Soriano-Ayala et al. (2021)	1	25 & 22	2x5	V&P; Off&On	-	.503-.772	-	-	**V = .75-.87 P = .68-.85**
VADRI-MX	Aizpitarte et al. (2019)	1	19	2x3	V&P; Off&On	**P = .92 V = .94**	**.81-.93**	-	**V = .957 P = .973**	**V = .800-.980 P = .747-.973**
EMVN	García-Carpintero et al. (2018)	1	32	9	V&P; Off&On	-	V = .771-.865 P = .611-.888	-	**V = .962 P = .966**	-
PMWI-SF	Tolman (1999)	1	14	2	V; Off&On	**6 months = .942 life = .888**	**.865-.934**	-	**.910 (6months)**	**.770-.885**
ESVIGA	Penado-Abilleiraa and Rodicio-García (2018)	1	13x2	2x5	V&P; Off&On	**.97**	**V = .929 P = .935**	-	**V = .933 P = .919**	V = .622-.849 P = .583-.800
ECS	Muñoz-Rivas et al. (2017)	1	2x5	2	V&P; Off	V = .51 P = .62	-	-	**V = .687 P = .677**	-
DVQ-R	Rodríguez-Díaz et al. (2017)	1	20	5	V; Off	**.85**	.64-.75	-	**.911**	**.622-.744**
M-CTS	Straus (1979) Muñoz-Rivas et al. (2007)	2	2x18	2x4	V; Off	V = .82; P = .77	.306-.819	ω: V = .43-.81 P = .48-.80	**V = .881-.906; P = .837-.898**	V = .359-.829 P = .380-.797
CMN	Osorio et al. (2012)	1	57	5	V; Off	**.93**	.75-.93	-	-	-
CTS	Straus (1979)	1	9x2	2	V&P; Off	-	-	-	-	-
DVQ-VP	Rodríguez-Franco et al. (2022)	1	2x20	2X10	V&P; Off	-	-	ω**: V = .81-.91 P = .78-.92**	**V = .964 P = .953**	**V = .796-.897 P = .709-.893**
VADRI	Javier-Juárez et al. (2022)	1	19	2	V; Off	**V = .96**	-	ω**: V = .84**	-	-

*When an 'x' is between two numbers, it signifies the multiplication of the number of items by the number of factors. If an '&' is included, it means that each factor has a different number of items. The values in bold are those that achieved the appropriate values. V = victimisation; P = perpetration; Off = Offline/face-to-face violence; On = Online or cyberviolence

Table 4. Scale, validity indicators

Scale	AFE	AFC	RMSEA	TLI	SRMR	GFI	CFI	x2/gl	AVE	AVE dimensions
CADRI	N	Y	**V = 0.015-0.041 P = 0.015-0.025**	-	**V = .045 (1) P = .049**	Ver: P = .907-.979 Conf: V = .913-.968	**G = .963 (1); P = .94 V = .936 (1)**	**G = 2.67 (1); P = 1.41-1.92 V = 1.47-1.87 (2)**	V = 0.176-0.297; P = 0.211-0.299;	V = 0.094-0.325; P = 0.005-0.410
DVQ	N	Y	G = 0.023-0.064; **F = 0.022 M = 0.026**	**G = 0.95-0.968 F = 9.78 M = .970**	-	-	G = .864-.971 (3); F = .970 M = .972 (1)	**G = 1.44-4.53 (3); F = 1.27 M = 1.2**	**0.586-0.592**	0.413-0.753
MSD V 2.0	N	Y	V = 0.053 P = 0.053	**V = 0.934 P = 0.909**	-	-	V = .946 P = .926	-	V = 0.484 P = 0.394	**V = .641-8.42 P = 0.600-0.801**
CARPA	Y	Y	**0.038**	-	-	-	**.99**	**2.27**	**V = .64; P = .58**	V = .53-.70; P = .43-.77
TDV-VP	Y	Y	V = 0.073 P = 0.066	V = 0.914 P = 0.877	-	-	V = .932 P = .900	**V = 3.228 P = 2.809**	-	**P = 0.31-0.550; V = 0.500-0.648**
VADRI-MX	Y	Y	**V = 0.05;** P = **0.03**	**V = 0.90; P = 0.93**	-	-	V = .; = .94	-	V = .542; P = .489	V = .447-.608; P = .374-.564
EMVN	Y	Y	V = 0.068; P = 0.068	V = 0.811; P = 0.832	-	-	V = .838; P = .857	**V = 3.08; P = 3.04**	V = 0.451; P = 0.483	-
PMWI-SF	N	Y	0.077	-	-	-	**.942**	**>5**	0.433	.337-.528
ESVIGA	Y	Y	V = 0.071 P.067	-	-	**V = .950; P = .953**	**V = .964; P = .971**	-	V = 0.488; P = 0.441	V = 0.267-0.653; P = 0.287-0.508
ECS	Y	Y	**V = 0.02; P = 0.03**	-	-	V = .1; P = .99	-	**V = 1.68; P = 2.30**	V = 0.352; P = 0.310	-
DVQ-R	Y	Y	**0.018**	**.94 (rob)**	-	-	**.95(robust)**	-	0.365	.293-.424
M-CTS	Y (1)	Y	V = 0.024-0.049; P = 0.024-0.029	-	**V = .049 P = .043 (1)**	**P = .962 V = .963** (1)	P = .675-.9 V = .91-.929	-	V = 0.307-0.382; P = 0.237+ 0.319	V = 0.210-0.383; P = 0.207-0.483
CMN	N	Y	0.05	**0.99**	-	-	**.99**	6.418	-	-
CTS	Y	Y	F = .066 M = .065	-	-	-	**F = .945 M = .979**	F = 5.97 M)5.79	-	-
DVQ-VP	N	Y	**V = .031; P = .028**	-	-	-	**V = .973; P = .967**	-	**V = 0.596; P = 0.524**	V = 0.496-0.687 P = 0.382-0.676
VADRI	N	Y	**V = .04**	**V = 0.98**	-	-	**V = .99**	**V = 1.70**	**-**	-

N = No; Y = yes; F = female; M = male; V = victimisation; P = perpetration; G = general scale. Values in bold are those that achieved the appropriate values

were found among them, 14 of the scales (87.5%) met the criteria for RMSEA, regarding the SRMR the two instruments that reported it obtained good values, all the scales reported the TLI except the EMVN (P = 0.832 V = 0.811); 4 of the 5 who presented the GFI (CADRI in one of the studies showed values of 0.907 for perpetration and 0.913 for victimisation). Regarding the CFI, although 15 of the 16 scales presented values greater than .90 in at least one of the validation studies, only six met the criteria (CADRI, CARPA, ESVIGA, DVQ-R, CMN, VADRI and DVQ-VP) in the case of the DVQ (English translated version), the 2016 study presented lower value (0.864).

Based on what was previously analysed, it could be concluded that the scale that presents the best and greatest evidence of construct validity is the CADRI which did not present TLI values and obtained scores above .90 for GFI. Other scales with acceptable or optimal values for all the indicators reported were the DVQ-VP, DVQ-R, VADRI and CARPA; however, it is worth mentioning that they did not report all the indicators and

those indicators with unacceptable values could have been suppressed.

Regarding the convergent and discriminant validity, using the values of the factor loadings provided by the different articles, the Average Variance Extracted (AVE) was calculated for the global scales and for each of their factors. with the exception of the validation of the TDV-VP where reference was made to this indicator while other studies evaluated the relationship with depression, anxiety and stress (MSDV 2.0 and CARPA), the correlation with other violence scales (CDAQ, ECS, M-CTS, PMWI-SF, ESVIGA) or with fear, perception of abuse, perceived relationship quality, and attachment-related anxiety scales (DVQ). Analysing the AVE, on the global scale or divided according to whether the violence was experienced or perpetrated, the scales that reported values greater than .50 were the DVQ-VP, CARPA and DVQ. However, at the factor level, none of the previous scales had acceptable values, only the MSDV 2.0 and TDV-VP.

Only the TDV-VP article reported that the scale did not have discriminant validity, and in the 2015 DVQ study it is estimated from Spearman's correlation with the EPQ sincerity scale. None of the studies presented values for HTMT or HTMT2.

The invariance was only studied by two of the investigations. In the case of the DVQ-R it is mentioned that there is invariance for sex, while the DVQ-VP presents configural, metric and scalar invariance for sex.

Due to study variability, only those utilising Cronbach's alpha and AVE for perpetrator or victimisation were included in the meta-analysis. Table 5 summarises their descriptive analysis. Twelve studies addressed victimisation, with an average sample size of 1,537.25 (*SD* = 1566.2025). Participants mean age was 18.271 years (*SD* = 2.765), with an average item count of 24.5. Additionally, the Cronbach's Alpha coefficient for the overall scale generates a mean of .891. However, when considering the composite reliability calculated based on the loadings reported in the studies, the mean was higher (.924) with an Average Variance Extracted mean of .0.453; an .0442 RMSEA mean and 0.943 CFI mean.

Regarding perpetration of violence, only eight studies included the required indicators, with an average sample size of 1,087.5 (*SD* = 581.9558). The mean age was 16.9286 years (*SD* = 2.03), with an average of 20.375 items. In terms of Cronbach's Alpha, the mean value was .859, and the Composite Reliability (CR) was .919. The average value of the Average Variance Extracted (AVE) was .398, the Root Mean Square Error of Approximation (RMSEA), reported by only seven studies, had a mean of .03 and the CFI a .919 (Table 5).

The reliability meta-analysis was performed in two parts. First, the investigations whose validations included instruments that contemplated the violence perpetrated (n = 8) were reviewed, later those that took into account the violence experienced (n = 12) that included a calculation of Cronbach's alpha and successively the same with AVE as moderator with violence perpetrated and violence experienced (Table 6).

The meta-analysis revealed high heterogeneity in perpetrated dating violence instrument reliability, explaining 99.51% of variance. H^2 indicates variability largely stems from genuine study differences. Cronbach's Alpha (α) exceeded .70 in most cases. Similarly, experienced violence instruments showed 99.39%

Table 5. Measures summarise the sample, mean age, initial and final scale items, and Cronbach's Alpha of the articles reviewed

Criteria	*n*	Missing	Mean	Median	*SD*	Minimum	Maximum
Experiencied							
Sample	12	0	1537.25	1015.5	1566.203	447	6138
Age	11	1	18.271	18.5	2.765	14.19	23.47
Total items	12	0	24.5	18.5	13.628	12	57
Cronbach's Alpha	12	0	0.891	0.894	0.045	0.82	0.96
Composite reliability	12	0	0.924	0.93	0.045	0.831	0.981
AVE	12	0	0.453	0.468	0.137	0.211	0.64
RMSEA	12	0	0.044	0.046	0.021	0.015	0.077
CFI	11	1	0.943	0.946	0.042	0.838	0.99
Perpetrated							
Sample	8	0	1087.5	970.5	581.953	477	2058
Age	7	1	16.929	16.15	2.03	14.19	20.1
Total items	8	0	20.375	18.5	7.23	12	32
Cronbach's Alpha	8	0	0.859	0.88	0.085	0.73	0.95
Composite reliability	8	0	0.919	0.922	0.062	0.783	0.973
AVE	8	0	0.397	0.418	0.129	0.176	0.58
RMSEA	7	1	0.040	0.03	0.022	0.015	0.068
CFI	6	2	0.919	0.923	0.039	0.857	0.971

Table 6. Heterogeneity Statistics

	AVE as moderator	Tau	Tau²	I²	H²	R²	*df*	Q	*p*
Perpetrated	No	0.107	0.0114 (*SE* = 0.0061)	99.51%	204.473	-	7	1630.802	< .001
	Yes	0.048	0.0023 (*SE* = 0.0014)	97.53%	40.447	79.75%	7	223.358	< .001
Experienced	No	0.076	0.0058 (*SE* = 0.0024)	99.39%	163.412	-	12	2325.452	< .001
	Yes	0.037	0.0014 (*SE* = 6 e-04)	97.25%	36.332	71.56%	11	353.733	< .001

heterogeneity, with reliability surpassing .75 in most studies (*n* = 8). Using the Hakstian-Whalen transformation, the average adjusted internal consistency for perpetrated violence scales was 0.499 (CI: 0.425-0.573), compared to an unadjusted 0.859 (CI: 0.801-0.918), likely influenced by heterogeneity. For victimisation scales, values dropped from .891 (CI: .866-.917) to .531 (CI: .492-.571) after adjustment (Figure 2).

Discussion

Following the review, it can be concluded that there are significant differences in the concepts, perspectives, methodologies, analytical strategies and indicators used in the evaluation of dating violence, even when the same instrument is used, a fact that makes comparisons less viable. None of the studies met all the current criteria of reliability and validity, as was the case in the Systematic Review by Tarriño-Concejero et al. (2023), where the use of the COSMIN guidelines was used as a criterion.

Concerning reliability, some scales provided values by dimension or factor, while others differentiated between violence perpetrated and experienced, and others even for the overall scale. The lowest global values were reported by the ECS and the M-CTS. At the subscale (perpetrated or experienced) or dimension/factor level, values were reported for the CADRI, TDV-VP, M-CTS, EMWI-SF, DVQ-R, and DVQ. Using more appropriate indicators, the lowest values were reported for the M-CTS with Omega.

Measurement reliability is crucial and requires careful consideration. In this regard, Cronbach's Alpha, widely used for continuous data, is unsuitable for categorical data such as Likert scales due to skewed distributions (Peterson & Kim, 2013). Considering this, we advocate for the adoption of more suitable measures tailored to effectively address the intricacies of categorical data. These include ordinal alpha, McDonald's omega, Raykov's Omega, or composite reliability (CR), all explicitly designed to navigate the nuances of categorical data. Among reviewed studies, the CARPA scale demonstrated strong reliability, but only six studies employed these suitable indicators, highlighting the need for greater methodological consensus (Cook & Beckman, 2006; Hayes & Coutts, 2020).

While these six studies demonstrated statistical rigour in indicator selection, their exclusion from our meta-analysis was due to the lack of uniformity and limited adoption of these approaches. Moving forward, we offer two recommendations. Firstly, researchers should consider supplementing their analyses with Cronbach's Alpha to ensure continuity and comparability with prior studies. Secondly, given the unique characteristics of the measurement context, we advocate for the use of tailored measures such as Omega or CR (Doval et al., 2023). This approach will enhance methodological rigour and promote harmonisation within categorical data research (Toro et al., 2022).

In evaluating validity within the reviewed studies, many neglected content validity. Instead, they often used pre-translated instruments or created custom tools, often neglecting pivotal cultural considerations (Arafat et al., 2016; Pedrosa et al., 2014). Furthermore, when it came to evaluating construct validity, the diversity in the model fit indices utilised precluded the comprehensive analysis of this criterion within the scope of our meta-analysis. While direct inter-study comparisons were not feasible, an assessment of the indicators presented across the diverse investigations underscores the commendable model fit and overall parsimony demonstrated by the CADRI, VADRI-MX, M-CTS, DVQ-R, CMN, and DVQ-VP.

Regarding convergent validity, many scales analysed it through comparisons with other measurement instruments. In the case of the MSD V 2.0, one study employed the self-perception of health item, and in one study of the DVQ, EPQ personality traits were used for this purpose. The only study that reported the Average Variance Extracted (AVE) was the validation of the TDV-VP. While all articles with factor loading values included AVE calculations, discrepancies in decimal places (ranging from one to three) could compromise precision (Benitez Muñoz & Muñoz Bandera, 2014; López-Barranco et al., 2022a). Given that the AVE offers insights into scale measurement quality by explaining variance relative to measurement error and facilitates discriminant validity assessment, its inclusion in future studies is crucial (Cheung et al., 2023; Henseler et al., 2015).

Discriminant validity was addressed in validation studies of the TDV-VP, EMVN, and DVQ, often through correlation analysis. Notably, none utilised modern cri-

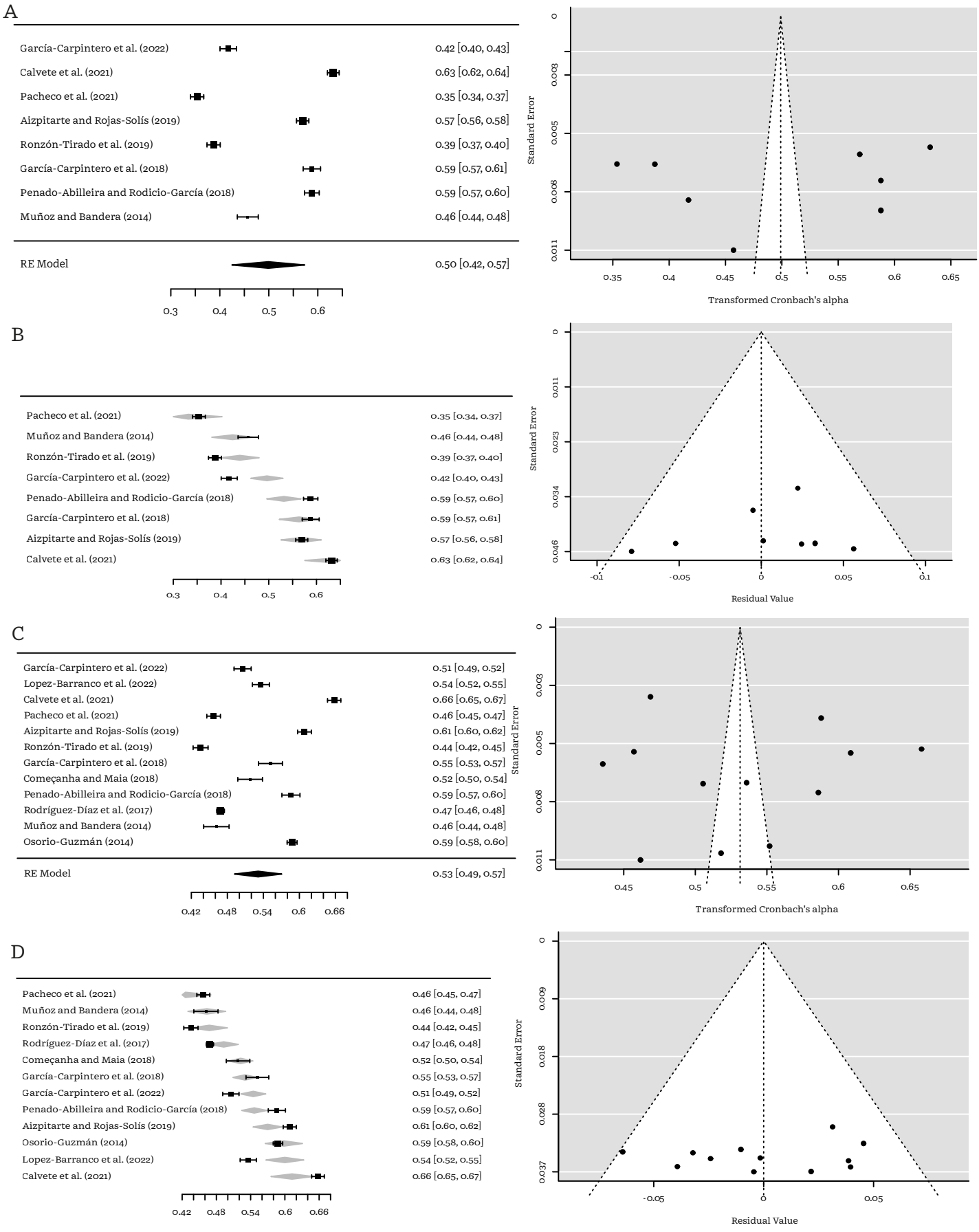

Figure 2. Meta-analysis of dating violence instruments A: Perpetration; B: Perpetration using AVE as moderator; C: Victimization and D: Victimization using AVE as moderator

teria such as HTMT or HTMT2. Concurrent validity was assessed in the DVQ, MSD V 2, and CARPA, comparing them with anxiety or depression tests such as the DAS-21, while the VADRI-MX, and M-CTS were compared with partner violence tests. The PMWI-SF used cutoff points, and the ESVIGA evaluated this validity through score correlations. Predictive validity was limited, with the DVQ using hierarchical multiple regression and the PMWI-SF relying on average scores across the lifespan. These findings highlight challenges with diverse measures, variable results, and a lack of longitudinal studies tracking violence cases (Henseler et al., 2015).

One noteworthy observation is the scarcity of information on measurement invariance, despite many studies making gender comparisons without verifying this aspect. It is worth highlighting that among the three scales that explore this aspect, only the DVQ-VP delves deeply into the matter. As cross-cultural and gender-focused violence research grows, future validations must prioritise measurement invariance to ensure robust and reliable results (Byrne, 2008; Raykov, 2004)

In this review, the most used scales in this review were the CADRI and DVQ (3 studies), with the latter adapted into the DVQ-R and DVQ-VP, which assesses perpetrated violence. Another frequently used scale was the M-CTS ($n = 2$). Notably, the latter three scales focus exclusively on face-to-face violence, highlighting the need to revise them to include cyber violence. Given the role of technology in modern relationships, especially among adolescents and young adults engaging on social media (Muñoz-Fernández et al., 2023; Thulin et al., 2023), incorporating cyber violence items is critical. Instruments such as the CARPA and VADRI-MX, addressing both online and offline violence, are promising for future evaluations.

Limitations

This research faced limitations, including the scarcity of publications from many Latin American countries in indexed journals and databases. Data inconsistency was another issue, with variations in presentation, such as decimal places or missing values. While efforts were made to standardise indicators such as CR and AVE using factor weights from articles, not all studies provided them or used standardisation, making complete calculations impossible. These inconsistencies likely stem from the broad time span of the studies and changes in APA guidelines. Furthermore, the lack of research on non-school populations and non-university-educated young adults remains a concern.

Funding

This work has been carried out with the collaboration of the Asociación Universitaria Iberoamericana de Postgrado (AUIP), European Regional Development Funds (European Union and Principality of Asturias) through the Science, Technology and Innovation Plan (AYUD/2021/51411) and the Spanish State Research Agency of the Ministry of Economic Affairs and Digital Transformation (MCI-21-PID2020-114736GB-100).

References

*references used in the Systematic Review study

*Aizpitarte, A., & Rojas-Solís, J. (2019). Factor structure of the violence in adolescents' dating relationships inventory for Mexican youth (VADRI-MX). *International Journal of Psychological Research, 12*(2), 29-36. https://doi.org/10.21500/20112084.4222

Alexander, E., Backes, B., & Johnson, M. (2022). Evaluating measures of intimate partner violence using Consensus-Based Standards of Validity. *Trauma, Violence & Abuse*, *23*(5), 1549-1567. https://doi.org/10.1177/15248380211013413

Arafat, S. M., Chowdhury, H. R., Qusar, M. S., & Hafez, M. S. (2016). Cross cultural adaptation & psychometric validation of research instruments: A methodological review. *Journal of Behavioral Health, 5*, 129-136.

*Benitez Muñoz, J., & Muñoz Bandera, J. (2014). Factor analysis of scores CADRI in Spanish college teenagers. *Universitas Psychologica*, *13*(1), 175-186. https://doi.org/10.11144/Javeriana.UPSY13-1.afpc

Byrne, B. (2008). Testing for multigroup equivalence of a measuring instrument: A walk through the process. *Psicothema*, *20*(4), 872-882. https://www.psicothema.com/pii?pii=3569

*Calvete, E., Fernández-González, L., Orue, I., Machimbarrena, M., & González-Cabrera, J. (2021). Validación de un cuestionario para evaluar el abuso en relaciones de pareja en adolescentes (CARPA), sus razones y las reacciones. *Revista de Psicología Clínica con Niños y Adolescentes*, *8*(1), 60-69. https://doi.org/10.21134/rpcna.2021.08.1.8

Centers for Disease Control and Prevention. (2016). *Understanding teen dating violence. Fact Sheet*. CDC.

Centers for Disease Control and Prevention. (27 de January de 2023). *Teen Dating Violence, Print*. Centers for Disease Control and Prevention https://acortar.link/5gjpXv

Cheung, G., Cooper-Thomas, H., Lau, R., & Wang, L. (2023). Reporting reliability, convergent and discriminant validity with structural equation modeling: A review and best-practice recommendations. *Asia Pacific Journal of Management*, *41*, 745–783 (2024). https://doi.org/10.1007/s10490-023-09871-y

Cho, G., Hwang, H., Sarstedt, M., & Ringle, Ch. M. (2020). Cutoff criteria for overall model fit indexes in generalizedstructured component analysis. *Journal of Marketing Analytics, 8*, 189-202. https://doi.org/10.1057/s41270-020-00089-1

*Começanha, R., & Maia, A. C. (2018). Screening tool for psychological intimate partner violence: Portuguese validation of the Psychological Maltreatment of Women Inventory. *Violence and Victims, 33*(1), 75-90. https://doi.org/10.1891/0886-6708.33.1.75

Cook, D. A., & Beckman, T. J. (2006). Current concepts in validity and reliability for psychometric instruments: Theory and application. *The American Journal of Medicine, 119*(2), 649-659. https://doi.org/10.1016/j.amjmed.2005.10.036

Doval, E., Viladrich, C., & Angulo-Brunet, A. (2023). Coefficient Alpha: The resistance of a classic. *Psicothema*, *35*(1), 5-20. https://doi.org/10.7334/psicothema2022.321

Exner-Cortens, D., Gill, L., & Eckenrode, J. (2016). Measurement of adolescent dating violence: A comprehensive review (Part 1, behaviors). *Aggression and Violent Behavior, 27*, 64-78. https://doi.org/10.1016/j.avb.2016.02.007

Fernet, M., Lapierre, A., & Cousineau, M. (2019). A systematic review of literature on cyber intimate partner victimization in adolescent girls and women. *Computers in Human Behavior, 100*, 11-25. https://doi.org/10.1016/j.chb.2019.06.005

Ferrando, P. J., Lorenzo-Seva, U., Hernández-Dorado, A., & Muñiz, J. (2022). Decálogo para el análisis factorial de los ítems de un test. *Psicothema*, *34*(1), 7-17. https://doi.org/10.7334/psicothema2021.456

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39-50. https://doi.org/10.2307/3151312

Gadermann, A., Guhn, M., & Zumbo, B. (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research, and Evaluation, 17*(3), 1-14. https://doi.org/10.7275/n560-j767

*Garcia-Carpintero, M., Rodriguez-Santero, J., & Maria Porcel, A. (2018). Diseño y validación de la escala para la detección de violencia en el noviazgo en jóvenes en la Universidad de Sevilla. *Gaceta Sanitaria, 32*(2), 121-128. https://doi.org/10.1016/j.gaceta.2017.09.006

*García-Carpintero-Muñoz, M., Tarriño-Concejero, L., Gil-García, E., Pórcel-Gálvez, A., & Barrientos-Trigo, S. (2022). Short version of the Multidimensional Scale of Dating Violence (MSDV 2.0) in Spanish-language: Instrument development and psychometric evaluation. *Journal of Advanced Nursing, 79*(4), 1610-1631. https://doi.org/10.1111/jan.15300

Gracia-Leiva, M., Puente-Martínez, A., Ubillos-Landa, S., & Páez-Rovira, D. (2019). La violencia en el noviazgo (VN): una revisión de meta-análisis. *Anales de Psicología*, *35*(2), 300-313. https://doi.org/10.6018/analesps.35.2.333101

Hakstian, A. R., & Whalen, T. E. (1976). A k-sample significance test for independent alpha coefficients. *Psychometrika, 41*, 219-231. https://doi.org/10.1007/BF02291840

Hayes, A., & Coutts, J. (2020). Use Omega Rather than Cronbach's Alpha for Estimating Reliability. But... *Communication Methods and Measures, 14*(1), 1-24. https://doi.org/10.1080/19312458.2020.1718629

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for metaanalysis.* Academic Press.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43,* 115-135. https://doi.org/10.1007/s11747-014-0403-8

Higgins, J. P., &Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*(11), 1539-1558. https://doi.org/10.1002/sim.1186

International Test Commission (ITC, 2017). *The ITC guidelines for translating and adapting tests* (2nd ed.). https://www.intestcom.org

*Javier-Juárez, P., Hidalgo-Rasmussen, C. A., Chávez-Flores, Y. V., Torres-Chávez, L., & Rosales-Damián, G. (2022). Relación entre el abuso cara a cara y digital en el noviazgo con la calidad de vida relacionada con la salud en adolescentes mexicanos. *Cadernos de Saúde Pública*, *38*(8), 1-12. https://doi.org/10.1590/0102-311xes071121

Jennings, W., Okeem, C., Piquero, A., Sellers, C., Theobald, D., & Farrington, D. (2017). Dating and intimate partner violence among young persons ages 15–30: Evidence from a systematic review. *Aggression and Violent Behavior, 33*, 107-125. https://doi.org/10.1016/j.avb.2017.01.007

Juarros-Basterretxea, J., Ocampo, N. Y., Herrero, J., & Rodríguez-Díaz, F. J. (2022). Dyadic analysis of emotional intimate partner violence: An estimation of dyadic patterns and influencing individual, relational, and couple factor. *The European Journal of Psychology Applied to Legal Context, 14*(2), 105-111. https://doi.org/10.5093/ejpalc2022a10

Kelley, K., & Pornprasertmanit, S. (2016). Confidence intervals for population reliability coefficients: Evaluation of methods, recommendations, and software for composite measures. *Psychological Methods, 21*(1), 69-92. https://doi.org/10.1037/a0040086

Lai, K. (2021). Fit difference between nonnested models given categorical data: Measures and estimation. *Structural Equation Modeling, 28*(1), 99-120. https://doi.org/10.1080/10705511.2020.1763802

*Lara, L., & Lopez-Cepero, J. (2021). Psychometric Properties of the Dating Violence Questionnaire: Reviewing the evidence in Chilean youths. *Journal of Interpersonal Violence, 36*(5-6), 2373-2392. https://doi.org/10.1177/0886260518760612

*Lopez-Barranco, P. J., Jimenez-Ruiz, I., Leal-Costa, C., Andina-Diaz, E., Lopez-Alonso, A., & Jimenez-Barbero, J. (2022a). Violence in dating relationships: Validation of the CADRI Questionnaire in a young adult population. *International Journal of Environmental Research and Public Health, 19*(17), 11083. https://doi.org/10.3390/ijerph191711083

López-Barranco, P. J., Jiménez-Ruiz, I., Pérez-Martínez, M. J., Ruiz-Penin, A., & Jiménez-Barbero, J. A. (2022b). Systematic review and meta-analysis of the violence in dating relationships in adolescents and young adults. *Revista Iberoamericana de Psicología y Salud, 13*(2), 73-84. https://doi.org/10.23923/j.rips.2022.02.055

*López-Cepero, J., Eduardo Fabelo, H., Rodríguez-Franco, L., & Rodríguez-Díaz, J. (2016). The Dating Violence Questionnaire: Validation of the Cuestionario de Violencia de Novios using a college sample from the United States. *Violence and Victims, 31*(3), 438-456. https://doi.org/10.1891/0886-6708.VV-D-14-00077

Love, H. A., Spencer, C. M., May, S. A., Mendez, M., & Stith, S. M. (2020). Perpetrator risk markers for intimate terrorism and situational couple violence: A meta-analysis. *Trauma, Violence & Abuse*, *21*(5), 922931. https://doi.org/10.1177/1524838018801331

Marcos, V., Cea, B., Novo, M., & Seijo, D. (2023). Contrasting cognitive competence of victimized youngsters in dating relations. *Revista Iberoamericana de Psicología y Salud, 14*(2), 68-74. https://doi.org/10.23923/j.rips.2023.02.067

Martínez-Soto, A., & Ibabe, I. (2022). Recommended Instruments for analyzing cyber dating violence: A systematic review. *The Spanish Journal of Psychology, 25*, 1-20. https://doi.org/10.1017/SJP.2021.50

Mokkink, L., De Vet, H., Prinsen, C., Patrick, D., Alonso, J., & Bouter, L. (2018). COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measure. *Quality of Life Research*, *27*(5), 1171-1179. https://doi.org/10.1007/s11136-017-1765-4

Muñoz-Fernández, N., Sánchez-Jiménez, V., Rodríguez-deArriba, M.-L., Nacimiento-Rodríguez, L., & Rosario del Rey, P. (2023). Traditional and cyber dating violence among adolescents: Profiles, prevalence, and short-term associations with peer violence. *Aggressive Behavior*, *49*(3), 261-273. https://doi.org/10.1002/ab.22069

*Muñoz-Rivas, M., Andreu Rodriguez, J., Graña Gomez, J., O'Leary, D., & González, M. (2007). Validación de la versión modificada de la Conflicts Tactics Scale (M-CTS) en población juvenil española. *Psicothema*, *19*(4), 693-698. https://www.psicothema.com/pdf/3418.pdf

*Muñoz-Rivas, M., Redondo, N., Zamarrón, M., & González, M. (2017). "Escala de coerción sexual (ECS) en relaciones de noviazgo: propiedades psicométricas en población juvenil española. *Behavioral Psychology*, *25*(3), 547-561.

*Osorio-Guzmán, M. (2014). Psychometric properties of the Dating Abuse Questionnaire in a sample of Mexican female students of middle and higher level. *Salud Pública de México, 56*(1), 40-47.

*Pacheco, J., Rey-Anacona, C., Mendez, J., & Briceno, M. (2021). Colombian Validation of the Conflict in Adolescent Dating Relationships Inventory (CADRI). *Revista Iberoamericana de Diagnóstico y Evaluación-E Avaliacao Psicologica, 58*(1), 141-152. https://doi.org/10.21865/RIDEP58.1.12

Page, M. J., McKenzie, J. E., Bossuyt, P. M. Boutron, I., Hoffmann, T. C. Mulrow, C. D., Shamseer, L., Tetzlaff, J. M. Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T. Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*(71). https://doi.org/10.1136/bmj.n71

Pedrosa, I., Suárez-Álvarez, J., & García-Cueto, E. (2014). Evidencias sobre la validez de contenido: avances teóricos y métodos para su estimación. *Acción Psicológica, 10*(2), 3-18. https://doi.org/10.5944/ap.10.2.11820

Peterson, R. A., & Kim, Y. (2013). On the relationship between coefficient alpha and composite reliability. *Journal of Applied Psychology, 98*(1), 194-198. https://doi.org/10.1037/a0030767

*Penado-Abilleira, M., & Rodicio-García, M. (2018). Development and validation of an Adolescent Gender-Based Violence Scale (ESVIGA). *Anuario de Psicología Jurídica*, *28*(1), 49-57. https://doi.org/10.5093/apj2018a10

*Presaghi, F., Manca, M., Rodríguez-Franco, L., & Curcio, G. (2015). A Questionnaire for the Assessment of Violent Behaviors in Young Couples: The Italian Version of Dating Violence Questionnaire (DVQ). *Plos One, 10*(5), e0126089. https://doi.org/10.1371/journal.pone.0126089

Prinsen, C. A. C., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., de Vet, H. C. W., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 27(5), 1147-1157. https://doi.org/10.1007/s11136-018-1798-3

Raykov, T. (1997). Scale reliability, Cronbach's coefficient Alpha, and violations of essential tau-equivalence with fixed congeneric components. *Multivariate Behavioral Research*, 32(4), 329-353. https://doi.org/10.1207/s15327906mbr3204_2

Raykov, T. (2004). Behavioral scale reliability and measurement invariance evaluation using latent variable modeling. *Behavior Therapy, 35*(2), 299-331. https://doi.org/10.1016/S0005-7894(04)80041-8

*Rodríguez-Díaz, F., Herrero, J., Rodríguez-Franco, L., Bringas-Molleda, C., Paíno-Quesada, S., & Pérez, B. (2017). Validation of Dating Violence Questionnarie-R (DVQ-R). *International Journal of Clinical and Health Psychology, 17*(1), 1-8. https://doi.org/10.1016/j.ijchp.2016.09.001

*Rodriguez-Franco, L., Juarros-Basterretxea, J., Paino-Quesada, S., Herrero, J., & Rodriguez-Diaz, F. (2022). Dating Violence Questionnaire for Victimization and Perpetration (DVQ-VP): An interdependence analysis of self-reports. *International Journal of Clinical and Health Psychology, 22*(1), 100276. https://doi.org/10.1016/j.ijchp.2021.100276

*Ronzón-Tirado, R., Muñoz-Rivas, M., Zamarròn Cassinello, M., & Redondo Rodríguez, N. (2019). Cultural Adaptation of the Modified Version of the Conflicts Tactics Scale (M-CTS) in Mexican adolescents. *Frontiers in Psychology, 10*, 619. https://doi.org/10.3389/fpsyg.2019.00619

Sánchez-Meca, J., Marín-Martinez, F., & López-López, J. (2013). Metodología del meta-análisis. En F. Sarabia Sánchez, *Métodos de investigación social y de la empresa* (págs. 447-470). Ediciones Pirámide.

*Soriano-Ayala, E., Sanabria-Vals, M., & Cala, V. (2021). Design and Validation of the Scale TDV-VP Teen Dating Violence: Victimisation and Perpetration [Violencia en parejas adolescentes: victimización y perpetración] for Spanish Speakers. *International Journal of Environmental Research and Public Health*, *18*(2), 421. https://doi.org/10.3390/ijerph18020421

Tarriño-Concejero, L., Gil-García, E., Barrientos-Trigo, S.-T., & García-Carpintero-Muñoz, M. (2023). Instruments used to measure dating violence: A systematic review of psychometric properties. *Journal of Advanced Nursing, 79*(4), 1267-1289. https://doi.org/10.1111/jan.15374

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

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

Thulin, E., Kernsmith, P., Fleming, P., Heinze, J., Temple, J., & Smith-Darden, J. (2023). Coercive-sexting: Predicting adolescent initial exposure to electronic coercive sexual dating violence. *Computers in Human Behavior, 141,* 107641. https://doi.org/10.1016/j.chb.2022.107641

Toro, R., Peña-Sarmiento, M., Avendaño-Prieto, B., Mejía-Vélez, S., & Bernal-Torres, A. (2022). Análisis empírico del coeficiente Alfa de Cronbach según opciones de respuesta, muestra y observaciones atípicas. *Revista Iberoamericana de Diagnóstico y Evaluación*, *63*(2), 17-30. https://doi.org/10.21865/RIDEP63.2.02

Torres, L., Morejón, Y., Medina, V., Urgilés, P., & Jiménez-Ruiz, I. (2022). Estudio de la prevalencia de las tácticas empleadas por los adolescentes ecuatorianos para la resolución de conflictos en el noviazgo. *Revista Iberoamericana de Psicología y Salud, 13*(1), 1-14. https://doi.org/10.23923/j.rips.2022.01.050

*Viejo, C., Sanchez, V., & Ortega-Ruiz, R. (2014). Physical dating violence: The potential understating value of a bi-factorial model. *Anales de Psicología, 30*(1), 171-179. https://doi.org/10.6018/analesps.30.1.141341

Xia, Y., & Yang, Y. (2019). RMSEA, CFI, and TLI in structural equation modeling with ordered categorWWical data: The story they tell depends on the estimation methods. *Behavior Research Methods, 51*(1), 409-428. https://doi.org/10.3758/s13428-018-1055-2

Yanez-Peñunuri, L., Hidalgo-Rasmussen, C. A., & Chávez-Flores, Y. V. (2019). Systematic review of dating violence questionnaires in Ibero-America and evaluation of their measurement properties. *Ciência & Saúde Coletiva, 24*, 2249-2262. https://doi.org/10.1590/1413-81232018246.19612017

Zumbo, B., Gadermann, A., & Zeisser, C. (2007). Ordinal versions of Coefficients Alpha and Theta for Likert Rating Scales. *Journal of Modern Applied Statistical Methods, 6*(1), 21-29. https://doi.org/10.22237/jmasm/1177992180