

# Investigating the reliability and validity of an intimate partner violence screening tool for use in physical therapy practice

## Abstract

We conducted a study to investigate the reliability and validity of a newly developed IPV screening tool. The literature shows that victims of intimate partner violence identify health-care practitioners as potential sources of support. Meanwhile health care professionals are often unprepared to identify and help victims appropriately. Our survey addressed psychosocial, socio-economic and physical constructs influencing prevalence and screening accuracy for IPV. Four experts in the fields of social sciences, education, and physical therapy gave feedback on the survey content validity. A rubric was designed to determine construct validity and inter-rater reliability. The new IPV screening tool was shown to have strong inter-rater reliability ( $ICC=0.71, p<0.001$ ), internal consistency (Cronbach's  $\alpha=0.80-1.0$  for all constructs,  $p<0.001$ ), construct validity, concurrent validity, and percent agreement (88.9-100%) among a expert panel of four qualified health professionals, including two licensed physical therapists and two licensed social workers with expertise in women's health issues.

**Keywords:** intimate partner violence, women's health, screening tool, validity, reliability

Volume 1 Issue 4 - 2017

Walton LM, Schbley BH, Milliner SW, Muvati FT, Zaeed N

Department of Physical Therapy, Andrews University, USA

**Correspondence:** Lori Maria Walton, Andrews University, Department of Physical Therapy, School of Health Sciences, Michigan, USA, Email [loriptdoc@gmail.com](mailto:loriptdoc@gmail.com)

**Received:** June 17, 2017 | **Published:** August 14, 2017

## Introduction

Currently, the Women Abuse Screening Tool (WAST), the Partner Violence Screening (PVS), and the Hit Insult Threaten Scream (HITS), are the most common IPV screening tools utilized by healthcare professionals. A study using 210 potentially eligible studies, 33 of which met inclusion criteria had the following results: "The most studied tools were the Hurt, Insult, Threaten, and Scream (HITS, sensitivity 30%-100%, specificity 86%-99%); the Woman Abuse Screening Tool (WAST, sensitivity 47%, specificity 96%); the Partner Violence Screen (PVS, sensitivity 35%-71%, specificity 80%-94%); and the Abuse Assessment Screen (AAS, sensitivity 93%-94%, specificity 55%-99%). Internal reliability (HITS, WAST); test-retest reliability (AAS); concurrent validity (HITS, WAST); discriminant validity (WAST); and predictive validity (PVS) were also assessed. Overall study quality was fair to good.<sup>1</sup>" The literature evaluating the reliability, validity, and comparing the tools' effectiveness is limited.

According to statistics published by the center for disease control (CDC), one in four women (24.3%) have been the subject of severe physical abuse by an intimate partner during a lifetime.<sup>2</sup> A study done by Rush College of Nursing reported that over a million women per year seek medical care for injuries caused by battering, yet only 10% of them are officially identified as victims of intimate partner violence (IPV). While women are routinely screened for problems such as health conditions and medical disorders, only one in ten are screened for intimate partner violence.<sup>3</sup> Several reasons for the low prevalence of IPV screening in the clinic were noted by different studies. The Rush College of Nursing study attributes lack of screening by clinicians to a lack of a comprehensive IPV screening tool. Current IPV screening tools have been criticized in studies for asking overly generalized questions that fail to incorporate context around which IPV arises. This makes them ineffective for use by PT's (Physical Therapists) in the clinic. Moreover, these current tools are not tailored specifically for use by PT's.

Evidence in another study suggests that therapists did not ask about IPV because they were concerned about misdiagnosing (3/22 13.6%).<sup>4</sup> Using a more thorough and construct specific screening tool may significantly decrease therapists' chances of misdiagnosing and hence alleviate their fear to perform the necessary IPV screens.

Barriers to IPV screening do not come from the therapists' end alone. An effective screening tool should account for client-caused factors that interfere with screening. One such client-caused factor is the client's perception of what abuse is. A client experiencing abuse, or one who grew up around IPV may view it less negatively. This increases the likelihood for such a client to respond in ways that would make it difficult to detect IPV. Our goal was to curb this problem by formulating survey questions around constructs that influence the victim's perspective of IPV. Literature from several studies<sup>5-8</sup> suggests that considering context, circumstances and motivating factors is key in investigation of intimate partner violence. Using evidence from the literature, the goal of our study was to create a screening tool that would address the different factors commonly associated with IPV. We considered the main socio-economic status (education, resource disparities, media and job stability) and psycho-emotional constructs (self-esteem, learned behavior, and religious beliefs) that influence how victims respond to a screening survey. Our newly developed IPV screening tool was created to address the needs of the physical therapist and in coordination with the musculoskeletal rating of a pain often seen in the rehabilitation setting.

## Methods

### Development of survey

The primary researcher and two other researchers created survey questions using information gleaned from literature on intimate partner violence screening. The created screening tool was given to four experts in the field and was evaluated using a rubric created by the authors that specifically addresses the constructs clarity, wordiness,

overlapping responses, appropriateness of responses, relationship to problem, psychosocial, physical, and socioeconomic. Feedback from experts indicated the need for changes regarding additional questions addressing “religious beliefs”. The screening tool was then sent back to the four experts. Changes were made, and an improved screening tool was sent out a second time with a rubric for feedback. Once the final evaluations from the experts was received it was analyzed statistically by the primary investigator. The main research question for this study, “Is the novel survey designed to detect IPV in physical therapy patients valid, reliable, and useful for clinical application?”.

The purpose of the proposed validity and reliability study was to create a survey designed to evaluate the presence of IPV to be used in the Physical Therapy clinic that would be valid, reliable, and useful for clinical application. The study tested the instrument for construct validity, content validity, internal consistency, and inter-rater reliability. Validity can be defined as whether the tool measures what it says it will measure. Content validity can be explained as whether the survey covers all of the necessary content found in the literature. Lastly, internal consistency is measured by whether the questions in the survey that are supposed to measure the same construct produce similar scores. After a thorough search of the literature, a survey was created using eight constructs, including: clarity, wordiness, overlapping responses, appropriateness of responses, relationships to problem, psychosocial, physical, and socioeconomic. A rubric was created to allow for experts to comment on how well the survey addressed the eight constructs. Our survey was created using research results from our literature review. A close review of the literature led to the need for socio-economic and emotional-psychological constructs in an IPV screening tool. Level of education, resource disparities, media, and job stability were addressed within the socio-economic category. The secondly category - emotional-psychological constructs - included self-esteem, spirituality, child abuse, learned behavior, and religious beliefs. The third construct category titled “physical constructs” including, topics of physical abuse, harm and domestic violence was also addressed in the survey.

The survey includes nineteen questions. Each question was formulated on a Likert scale, giving multiple options for each question. Two questions are in a visual analog scale format asking

the patient to rate their pain. A rubric was also created in order to provide subjects a formulated way to grade the survey. The rubric was set up also in a Likert scale format. Each construct was graded on a scale from zero to four on how well it was addressed in the survey. A score of zero indicates the absence of a construct. A score of “one” indicates an identified and measureable construct, but it was deemed “unacceptable”. A score of “two” indicates that the construct was “below expectations”. A score of “three” indicates the construct “meets expectations”. Lastly, a score of “four” would conclude that the construct “exceeded expectations”. An open comment box was added to the rubric for additional feedback. Statistical analysis was done using the Intraclass Correlation Coefficient (ICC) addressing reliability, Kappa Percent Agreement among experts, and Cronbach Alpha addressing internal consistency of each construct. Data from the rubric was evaluated using IBM SPSS version 22.0 statistic software by the primary investigator.

## Results

### Content validity

Experts in the field were selected for review of the newly developed screening tool: (1) A Women’s Health Physical Therapist, (2) Licensed Social Worker, (3) Professor of Social Work, and (4) a Physical Therapist Faculty with experience in Womens Health. The created survey and rubric was sent to the subjects for review via email. The subjects reviewed the survey, filled out the rubric and returned the results. Changes were made to the survey and the rubric after feedback was given from the experts. Changes were made addressing the consistent use of the term “intimate partner”, the order of the questions in the survey, removal of reference sources after each question, and removal of the heading for each section. A visual analog scale was added for assessing patient pain. Changes were also made to the rubric in order to make questions more specific. A table was added to the bottom of the rubric so that experts could check how well the survey addressed specific constructs from the survey questions. The survey and rubric were sent a second time to experts for final feedback. Experts were given a rubric to grade whether the eight constructs were addressed accurately and thoroughly in the survey Table 1.

**Table 1** ICC, Cronbach Alpha, and Percent Agreement between Experts values listed

Question	Cronbach alpha	% Agreement between experts rated as 3 or >
Clarity	0.875	91.70%
Wordiness	0.789	85.50%
Overlapping response	0.938	91.70%
Appropriateness of responses	0.875	91.70%
Relationship to problem	1	93.80%
Psychosocial	0.979	73.9% -75.9% with question 12 taken out
Socio-economic	0.911	82.80%
Physical	0.979	69.8% - 2.2% with question 11 taken out

### Construct validity

The Cronbach’s alpha (internal consistency) values range from .79-1.0 and were evaluated through analysis of percent agreement values,

which represented scores of 3 or greater. With the values for percent agreement ranged from 88.9%-100%, with an ICC value at .71. This combination provides excellent foundation for overall strong construct validity for each of the questions rated by our experts.

## Kappa statistic & ICC

SPSS was used to run the results of each reviewer's responses to each construct. The percent agreement between the responses was found to be 100% with the exception of the following constructs: Limitation of redundancy=97.7%; physical abuse construct=66.7% for all items calculated and "physical abuse construct" 88.9% with questions 8 and 9 deleted. The ICC for the entire survey rubric was .71( $p<.0005$ ).

## Reliability (internal consistency)

Cronbach's Alpha was calculated to measure the internal consistency (IC) of each construct that had been measured on the created IPV screening tool. Questions of the rubric were grouped into the following categories: Good IC ( $\alpha=.75-1$ ); Acceptable ( $\alpha=.5-.7$ ) Poor ( $\alpha<.5$ ). Cronbach alpha values were as follows: Psycho-social construct=.979; Socio-economic construct=.911; Physical abuse construct=.979; Relationship to IPV issue=1.0; Appropriateness of responses =.875; Limit redundancy=.938; Wordiness=.789; Clarity=.875. All Cronbach's alpha values fell within the "strongly related" category, supporting a strong internal consistency of constructs assessed by the tool. Figure 1 below shows Cronbach's alpha values for each of the constructs representing the internal consistency for each construct. The inter-rater reliability was calculated to determine the correlation/agreement between raters on responses using SPSS. All 46 questions of the rubric were evaluated at one time, revealing reliability between reviewers and an ICC =.71. The interclass correlation (ICC) for the survey between experts was considered good and acceptable for utilization in the clinic (Figure 2).

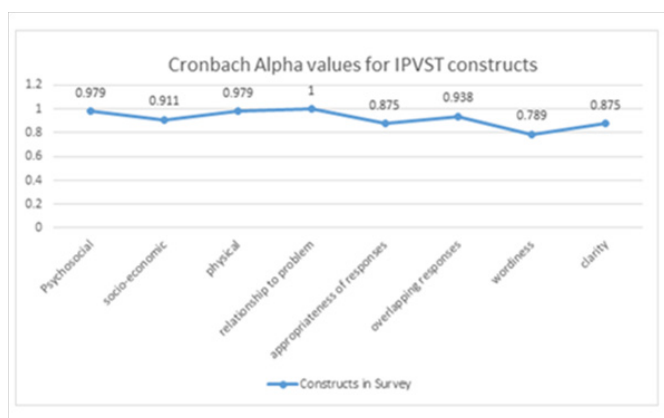


Figure 1 ICC (Inter-rater Reliability).



Figure 2 Construct validity.

Table 2 shows experts' rating of each question for clarity, wordiness, overlapping responses, appropriateness of responses and relationship to problem being addressed for each question. Each expert graded each question on a scale of 0-4.

## Discussion

Our study found the newly developed IPV Screening Tool for

Physical Therapists to hold a strong inter-rater reliability between experts, strong internal consistency, and percent agreement. Considering the strong internal consistency (.79-1.0 for all constructs), good intra-rater reliability (ICC=.71), and strong kappa scores for all constructs provides evidence of good to excellent construct validity for this screening tool. The combination of this information suggests the new IPV screening tool to be both valid (content and construct validity) and reliable (internal consistency).

Table 2 shows experts' rating of each question for clarity, wordiness, overlapping responses

0	Construct not identified or measured, "absent"
1	Construct identified and measured but "not acceptable"
2	Construct identified and measured but "below expectations"
3	Construct identified clearly and measured "meets expectations"
4	Construct clearly identified and measured more than once "exceeds expectations"

Physical question number 11 was found to have poor agreement between experts. This question was given a rating of <2 correlating with "below expectations" by two experts and a rating of 0 meaning this construct was "absent". This question addressed the topic of rape. Question 3 in the survey states, "If you have experienced verbal or physical abuse by your intimate partner what percentage of the pain are you currently experiencing and do you perceive to be related to your relationship with your intimate partner?" and patients are asked to fill out a likert scale. We will change the question to include sexual abuse (i.e. rape). Psychosocial question 12 was also found to have poor agreement between experts. This question was given >2 or "below expectations" by two experts and a >3 or "meets expectations" by one expert. Psychosocial question 12 is dealing with the topic of blackmailing. Question number 13 states, "Has your partner ever threatened to expose your private info via social media"? We will address the topic of blackmailing by changing the question to ask, "Has your partner ever threatened to expose your private info via social media if you don't do what they say"? With these changes, we believe we have addressed all inconsistencies determined by expert reviewers.

## Compare and contrast

Comparable studies were conducted at Rush University, which investigated the use of a comprehensive abuse-screening tool in an increased identification rate of IPV.<sup>3</sup> A sample population of 438 women ranging in age from 18 to 78 years participated. This study used factor analysis and examination of content by experts in the field, which is very similar to the methodology used in this study. Internal consistency, content validity, and discriminant validity were the specific psychometric properties tested. Data analysis results showed the WAST to be highly reliable with a Cronbach's alpha =.93 for physical abuse and .91 for verbal/emotional abuse. These Cronbach's alpha values were comparable to those of our study that had a Cronbach's alpha=.98 for physical abuse and .98 for psychosocial abuse.

To further Cronbach's alpha comparison, we used Nyberg's study that was conducted with randomly selected male participants (2013). This study used the Violence against Women Instrument (VAWI). According to the study, women's and men's exposure to IPV differs in

certain aspects because of the etiologies of the violence. Aspects such as motives for inflicting abuse, attitudes towards violence, context in which violence occurs and the consequences of inflicting abuse. For example, in assessing IPV differences between men and women, the study found men's abuse against women to be more severe, controlling and threatening. Cronbach's alpha values for this study

were .74 (psychological); .86 (physical); .82 sexual; and .88 total score. For all the constructs in Nyberg's study that were comparable to our constructs, our created tool had Cronbach's alpha values: .979 (psychological); .979 (physical including sexual) and a total score of .92. Table 3, contrasts the Cronbach's alpha results of coinciding constructs in this study compared to ours.

**Table 3** Contrasts the Cronbach's alpha results of coinciding constructs in this study compared to ours

Construct	Chronbach alpha for VAWI tool with male subjects	Chronbach alpha for our gender neutral created tool
Psychological/Psychosocial	0.74	0.979
Physical	0.86	0.979
Sexual	0.82	
Total score	0.88	

Of note, in addition to serving as a comparison for Cronbach alpha, the study discussed above concluded that a need exists for research instruments that assess IPV separately in male and female samples in order to ensure their suitability for the respective groups. Our created screening tool deviates from such a conclusion because it is gender neutral and refers to the abusive partner using words like, "significant other", "intimate partner" as opposed to more specific "husband", "wife" or "girlfriend".

In a cross-sectional study, Chen et al tested the reliability and validity of a brief 4-question instrument, the Hurt Insult Threaten Scream (HITS) among predominantly Hispanic women. Two hundred and two women completed the HITS and two other previously validated tools, the Index of Spouse Abuse-Physical Scale (ISA-P) and the Woman Abuse Screening Tool (WAST). Instruments were prepared in English and translated to Spanish. Reliability and validity of HITS were compared with the ISA-P and WAST. Performance measures of HITS were compared with the ISA-P or WAST as a criterion standard. Cronbach's alpha values were 0.76 and 0.61 respectively for the English and Spanish versions of HITS. When administered first and analyzed alone, the Spanish version of HITS had a reliability of 0.71. Compared to our study, the Spanish version of HITS had a similar reliability score (0.71) to the novel screening tool when used alone and a lower reliability score (0.61) when administered together with the version to Spanish women.

For Kappa statistic comparison we looked at two studies: (1) the Laurie studies. Ernst et al.<sup>9,10</sup> created a screening tool that they called the ongoing abuse Screen (OAS), which they compared to a previously validated tool called Abuse Assessment Screen (AAS). The intended advantage of the Ernst created tool over the AAS was the addition of a construct addressing currently ongoing IPV. The AAS and all other prior studies asked about IPV "within a year" only. The Kappa statistic in Ernst et al's study was found to be 0.28 for both AAS and OAS. When compared to our created tool, both the AAS and OAS had a much lower percent agreement than the 72.2-92% we calculated for our tool.

Laurie<sup>11</sup> investigated the validity of the Brief Inpatient Screen (BIS) for IPV among adult women. The Brief Inpatient Screen was designed to assess recent emotional, physical, and sexual abuse in a general inpatient medical-surgical setting. The study compared the BIS to the Composite Abuse Scale (CAS). Performance of the BIS was compared when used as a verbal versus when it was used as a written

screen. The kappa statistic or percent agreement for BIS was found to be 68% which falls in the moderate category. Based on findings of this study, the BIS has a lower percent agreement than our created IPVST (Kappa 72.2-90%) which reflected strong agreement and hence strong reliability compared to 88%-100% Kappa agreement for the current novel IPV Screening tool in this study.

This newly developed IPV screening tool reliability and validity study will serve as a pilot study for future use as a screening tool for Physical Therapists and other Allied Health Professionals. A study by McMillan et al.<sup>12</sup> reported, "Insufficient evidence exists to recommend for or against universal screening (for IPV). The rationale for this position is based on limitations of existing screening approaches including validation studies, insufficient evidence regarding effectiveness of services to which women can be referred once identified, a paucity of evidence that IPV screening improves the likelihood of positive health outcomes, and a lack of studies evaluating the potential harm associated with IPV screening." Reports from this study examine women who disclosed IPV within one year of the study, with interviews at baseline and every 6 months regarding IPV re-exposure and quality of life. Of the screened women 43% of participants dropped out or failed to be followed up. 41% of those not screened dropped out. At 18 months, recurrence of IPV screened vs nonscreened was 46%vs 53%. Screened vs non-screend women showed 0.2 standard deviation greater improvement in quality of life. This study concluded that its results did not produce sufficient evidence to support IPV screening. The lack of significant benefits for screening in McMillan et al.<sup>12</sup> study may have been limited by poor availability of services once IPV was identified, as well as a high dropout rate.

## Limitations and delimitations

One expert, who provided consistently low scores when compared to the other three expert reviewers, may have negatively skewed the results. The two questions ranking below a .50 include physical construct 11 dealing with rape, and psychosocial construct 12 dealing with blackmailing. All researchers agreed that these two constructs could have been better addressed in the survey. These two constructs are important for proper screening of IPV and were addressed by changing the questions mentioned above to include the proper wording to addresses the constructs to more equally represent the issues of "rape" and "blackmail".



## Future research suggestions

The main purpose of this research was to design and validate a survey tool with the intent of using the survey in the physical therapy clinic setting. Future research should focus on a pilot test for utilization in the clinic to determine clinical usefulness and provide further feedback from both patients and clinicians.

## Application of research

The purpose of the created survey was to enable Physical Therapists to screen for IPV to cover all constructs including: (1) physical, (2) socioeconomic, and (3) psychosocial constructs with relationship to physical pain often reported in physical therapy evaluation. The created survey can be used as a screening tool to inform care for those who are suffering from Intimate Partner Violence and musculoskeletal pain, concurrently. IPV is a real problem for 25% of women and is often neglected by the physical therapist because of lack of training and available screening tools. The purpose of creating this survey and providing validity and reliability testing of the tool is to promote utilization of this quick screening tool in the clinic for physical therapists and to inform physical therapy practice of micro and mesostructure violence that can impact a woman's health care.

## Conclusion

The new IPV screening survey was designed as an efficient and effective tool in the Physical Therapy clinic to promote screening for Intimate Partner Violence and inform the practice of physical therapy. The IPV screening tool was found to have excellent content and construct validity, internal consistency (Cronbach's  $\alpha = .80$  for all constructs), inter-rater reliability ( $ICC = .71$ ,  $p < .0005$ ) and percent agreement among a expert panel of four qualified health professionals, including two licensed physical therapists and two licensed social workers with expertise in women's health issues. Future research should focus on a pilot study for clinical application with feedback from both patients and clinicians regarding usefulness in the physical therapy setting for female patients.<sup>13–15</sup>

## Acknowledgements

None.

## Conflict of interest

Author declares there is no conflict of interest in composing this manuscript.

## References

1. Rabin RF, Jennings JM, Campbell JC, et al. Intimate Partner Violence Screening Tools: A Systematic Review. *Am J Prev Med*. 2009;36(5):439–445.
2. Black M, Basile K, Breiding M, et al. The national intimate partner and sexual violence survey (NISVS): 2010 Summary Report. National Center for Injury Prevention and Control, Centers for Disease Control and Prevention, Georgia; 2011.
3. Taylor WK. Development of the Women Abuse Screening Tool. Rush University College of Nursing; 1994.185 p.
4. Sprague S, Madden K, Simunovic N, et al. Barriers to screening for intimate partner violence. *Women Health*. 2012;52(6):587–605.
5. Belknap J. The invisible woman: Gender, crime, and justice, 2nd ed. Belmont, CA, USA; 2001.
6. Dasgupta SD. A framework for Understanding Women's Use of Nonlethal Violence in Intimate Hereosexual Relationships. *Violence Against Women*. 2002;8(11):1364.
7. Fleury SRE, Bybee D, Sullivan CM, et al. Contextual factors impacting battered women's intentions to reuse the criminal legal system. *Journal of Community Psychology*. 2006;34(3):327–342.
8. Praveen A, Naylor PB. Intimate partner violence: A narrative review of the biological and psychological explanations for its causation. *Aggression and Violent Behavior*. 2013;18:373–382.
9. AA Ernst, SJ Weiss, E Cham, et al. Comparison of three instruments for assessing ongoing intimate partner violence. *Med Sci Monit*. 2002;8(3):CR197–201.
10. Ernst AA, Weiss SJ, Cham E, et al. Detecting ongoing intimate partner violence in the emergency department using a simple 4-question screen: the OVAT. *Violence and victims*. 2004;19(3):375–384.
11. Laurie AR, Showalter J, Pratt T, et al. Validity of the Brief Inpatient Screen for intimate partner violence among adult women. *Women Health*. 2012;52(7):679–699.
12. MacMillan, Harriet L, Nadine WC, et al. Screening for intimate partner violence in health care settings: a randomized trial. *Jama*. 2009;302(5):493–501.
13. Ramsay J, Rutterford C, Gregory A, et al. Domestic violence: knowledge, attitudes and clinical practice of selected UK primary healthcare clinicians. *Br J Gen Pract*. 2012;62(602):647–655.
14. Nybergh L, Taft C, Krantz G. Psychometric properties of the WHO Violence Against Women instrument in a female population-based sample in Sweden: a cross-sectional survey. *BMJ open*. 2013;3(5):e002053.
15. PH Chen, S Rovi, M Vega, et al. Screening for domestic violence in a predominantly Hispanic clinical setting. *Fam Pract*. 2005;22(6):617–623.