Preprints.org

Article

Not peer-reviewed version

Developing a short-form Buss-Warren Aggression Questionnaire based on Machine Learning

Xiuyu Jiang , Junyi Li * , Yitian Yang

Posted Date: 2 August 2023

doi: 10.20944/preprints202308.0189.v1

Keywords: aggression questionnaire; machine learning; short-form questionnaire; adolescents

Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 August 2023 doi:10.20944/preprints202308.0189.v1

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Developing a Short-form Buss-Warren Aggression Questionnaire Based on Machine Learning

Xiuyu Jiang, Yitian Yang and Junyi Li *

College of Psychology; Sichuan Normal University; Chengdu(610066); China
* Correspondence: junyili@sicnu.edu.cn

Abstract: For adolescents, high aggression is often associated with suicide, physical injury, worse academic performance, and crime. Therefore, there is a need for early identification and intervention for highly aggressive adolescents. The Buss-Warren Aggression Questionnaire (BWAQ) consists of 34 items, and the longer the scale, the more likely participants are to make an insufficient effort response (IER), which reduces the credibility of the results and increases the cost of implementation. The study aimed to develop a shorter BWAQ using machine learning (ML) techniques to reduce the frequency of IER and decrease implementation costs meantime. First, an initial version of the short-form questionnaire was determined using Stepwise Regression and ANOVA F-test. Then, a machine learning algorithm determined the optimal short-form questionnaire (BWAQ-ML). Finally, the reliability and validity of the optimal short-form questionnaire were tested using independent samples. The BWAQ-ML has 88% fewer items than the BWAQ. It has AUC, accuracy, recall, precision, and F1 scores of 0.85, 0.85, 0.89, 0.83, and 0.86, respectively, and good psychometric properties. The BWAQ-ML can effectively measure individual aggression and can be used as a simplified version of BWAQ.

Keywords: aggression questionnaire; machine learning; short-form questionnaire; adolescents

1. Introduction

Aggressive behavior is prevalent in children and adolescents. In a survey of approximately 15,000 middle school students in China, the detection rate of highly aggressive adolescents was 23.5% [1]; in a survey of approximately 80,000 middle school students in the United States, 18.8% of adolescents reported committing aggression, and 20.1% of adolescents reported having been assaulted [2]. Additionally, highly aggressive individuals are more likely to maintain this high level of aggression. A short-term longitudinal study found 87.4% of adolescents were still highly aggressive one year later, compared to 52.3% of low-aggressive adolescents [3]. Aggressive behavior has many negative effects. High aggression can lead to physical injuries to others, with 700,000 adolescents (10-24 years old) in the United States alone being treated in hospitals for injuries from aggressive behavior in 2011 [4]. In addition, aggression is a potential risk factor for non-suicidal self-injury [5] and is positively associated with suicidal behavior [6,7]. In addition, highly aggressive adolescents are more likely to commit crimes [8], and academic performance is worse [9-11]. Therefore, efficient identification of highly aggressive adolescents and comprehensive interventions are essential for reducing social risk and maintaining mental health.

The Buss-Warren Aggression Questionnaire is one of the most widely used instruments for measuring and rating aggression, which was developed from the Buss-Perry Aggression Questionnaire (BPAQ) with a more concise and more transparent formulation of questions [12,13] and has good psychometric properties [6,14]. However, the BWAQ consists of 34 items, and participants must exert a certain degree of effort to complete all the items. In practice, participants do not answer every question carefully [15]. The incidence of insufficient effort response (IER) is up to 30% in all types of surveys [16]. As the length of the questionnaire increases, the proportion of participants seriously completing the questionnaire decreases, and the later the item, the worse the quality of their responses [17]. Therefore, the longer the questionnaire is, the more likely it is to result in more low-quality responses and lower questionnaire completion rates, making results less

© 2023 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 August 2023 doi:10.20944/preprints202308.0189.v1

accurate. Failure to deal with this appropriately can adversely affect the veracity of the study results [18]. An effective means of reducing IER is to reduce the length of the questionnaire, which is an ex-ante control method that reduces the perceived difficulty of the task for participants, thereby reducing the frequency of IER [16]. Although shorter versions of the Aggression Questionnaire exist, such as the BPAQ-12 and BWAQ-15 [19,20], both shorter versions of the questionnaire have more than ten items and still require some effort for participants to complete. In addition, organizing and completing a large-scale measurement will inevitably consume many human and material resources for teachers, clinicians, and researchers, which may hamper research efforts. Therefore, it is necessary to use appropriate methods to streamline the questionnaire as much as possible while ensuring its performance. A streamlined questionnaire reduces IER and research costs, improves measurement efficiency, and facilitates the researcher to measure more variables simultaneously in one administration.

A common method researchers use in adapting and simplifying psychometric instruments is exploratory factor analysis, but this method requires items to be retained for each dimension of the questionnaire. Therefore, making psychometric instruments more streamlined by traditional methods is challenging. In recent years, some researchers have applied machine learning techniques to the study of simplifying psychometric tools. Machine learning belongs to the field of artificial intelligence, which can use the patterns learned from training data in the prediction of new data, and it is a reliable data analysis technique that can make the questionnaire simplification research free from the constraints of dimensions. An increasing number of studies use Machine Learning techniques to simplify psychometric instruments, but they are mainly focused on clinical psychological assessment [21]. Wang reduced the Berg Balance Scale from 14 to 6 items based on Machine Learning techniques, with a 57% reduction in the number of items [22]. The R^2 of the short version of the questionnaire was more significant than 0.96, and the LoAs were less than 95%, indicating a good recognition performance. Lee used six Machine Learning algorithms to reduce the Insomnia Severity Index (contains seven items) and Epworth Sleepiness Scale (contains eight items), to a 6-item short-form questionnaire, with a 60% reduction in the number of items, and accuracy reached 0.93 [23]. Orrù reduced the Structured Inventory of Malingered Symptomatology (includes 75 items) to a 21-item short-form version with a 72% reduction in the number of items, retaining 92% of the variance of the original scale [24]. Morrison reduced the Cognitive Distortions Questionnaire (which involves 15 items) into a 5-item ultrashort version, which was reduced by 67% from the original questionnaire, with an R^2 of between 78.2 and 85.5% [25]. Lin reduced the Fugl-Meyer motor scale (which contains 50 items) into a short version of 10 items, a reduction of 80% in number, with Pearson's r between 0.88-0.98 with the original measurement tool [26]. Machine Learning techniques are a reliable method capable of simplifying psychometric instruments. However, past studies have primarily used R^2 as performance indicators and have not created the cutoff of short-form questionnaires. Machine Learning classification algorithms can provide performance indicators such as AUC, accuracy, recall, precision, and F1 Score, which can help researchers to assess the optimal cutoff for short-form scales.

Therefore, this study proposes simplifying the BWAQ using Machine Learning classification algorithms to develop a more streamlined short-form aggression questionnaire with an explicit cutoff to improve the efficiency of examining adolescent aggression and provide teachers, clinicians, and researchers with a more convenient tool.

2. Materials and Methods

2.1. Participants

Before the beginning of the survey, the investigator informed the participants that this study is to understand the current situation of adolescents, that there are no right or wrong answers, that the results of the survey will be used only for scientific research, that the questionnaire was collected anonymously, and that the data will be treated with strict confidentiality. The study was ethically approved by the Ethics Committee of Sichuan Normal University on 15 March 2023 (No.2023LS029).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 August 2023 doi:10.20944/preprints202308.0189.v1

The participants were 796 middle school students. The sample was divided into two parts: simplification samples, which were used to simplify the BWAQ, and validation samples, which were used to validate the short-form BWAQ. There were 340 middle school students in the simplification samples, with a mean age of 14.83±1.57 years. There were 200 female students, accounting for 58.8% of the simplification samples. Students categorized as highly aggressive by the BWAQ numbered 175 or 51.5%. Seventh, eighth, tenth, and eleventh graders comprised 84, 75, 113, and 68 students, or 24.7%, 22.1%, 33.2%, and 20%, respectively. Validation samples comprised 456 middle school students with a mean age of 15.3 ± 1.3 years. There were 265 females, or 58.1% of the validation sample, and 236 students, or 51.8% of the validation sample, whom the BWAQ-ML categorized as having highly aggressive tendencies. The percentages of grades 8 through 12 of the validation sample were 37.9%, 2.9%, 36.4%, 21.7%, and 1.1%, respectively.

2.2. Measurements

The BWAQ has 34 items belonging to the five dimensions: physical aggression, verbal aggression, hostility, anger, and indirect aggression. The scale was rated on a 5-point, with "1" indicating "not at all like me" and "5" indicating "completely like me", with higher scores indicating higher aggression. Maxwell revised the Chinese version of the BWAQ [27], based on which Zhang developed a norm of the BWAQ for students: 12 to 18 years old with total scores of 89 and above could be labeled as highly aggression [28,29]. The validity scale for this study used the Reactive-Proactive Aggression Questionnaire (RPQ) developed by Raine [30]. The RPQ is divided into two dimensions, reactive aggression (includes 11 items) and proactive aggression (includes 12 items). It is scored on a 3-point scale: 0 for never, 1 for sometimes, and 2 for often. Higher total RPQ scores indicate higher aggression. The RPQ has good psychometric properties in mainland China [31-33].

Carlo and Randall developed the Prosocial Tendencies Measure (PTM) to measure individuals' prosocial tendencies [34], and we used the PTM to test the discriminant validity of the short-form scale. PTM consists of 23 items and is scored on a 5-point, with "1" meaning "not at all like me" and "5" meaning "completely like me". Higher total PTM scores indicate higher prosocial tendencies. The PTM has good psychometric properties in mainland China [35].

2.3. Statistical Analysis

We analyzed data using Python 3.9 in PyCharm 2023.1.2 (Community Edition). The whole process is divided into two phases: simplification and validation. In the simplification phase, the simplification sample is randomly divided into a training set (80% simplification sample), used for feature selection and model training, and a test set (20% simplification sample) for evaluating the model's performance. The first step is Feature Selection. The initial version of the short-form BWAQ was determined by analyzing the training set using Stepwise Regression and ANOVA F-test to extract the most contributing items (i.e., the most important features) to the questionnaire results. The second step is Machine Learning modeling. The initial version of the short-form BWAQ were modeled using Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), and Naïve Bayes (NB), and the short-form BWAQ was evaluated according to AUC, accuracy, recall, and precision performance of the model to determine the best short-form questionnaire. The third step is to determine the cutoff. We calculate AUC, accuracy, recall, precision, and F1 scores based on the confusion matrix under different cutoff conditions to specify the cutoff of BWAQ-ML. During the validation phase, we verify the reliability and validity of BWAQ-ML using validation samples, and we test the reliability of BWAQ-ML as a streamlined version of BWAQ using RPQ and PTM.

3. Results

3.1. Simplifying BWAQ

We determined the initial version of the short-form questionnaire using Stepwise Regression and ANOVA F-test. At a significance level of 0.01, Stepwise Regression identified the eight most important features, which in descending order of importance, were items 29, 14, 21, 18, 7, 30, 1, and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 August 2023 doi:10.20944/preprints202308.0189.v1

23 of the BWAQ. The ANOVA F-test ranked the importance of the features differently than the Stepwise Regression (Table 1). In ANOVA F-test, the larger the F-value of a feature indicates its higher correlation with the outcome, and the p-value is the significance level of this correlation. Thirty-two out of all 34 items have a significance level below 0.001, and the top 25% are the most important features according to F-value in descending order, and they are items 29, 7, 21, 9, 12, 5, 33, and 14 of the BWAQ. Stepwise Regression and ANOVA F-test each identified eight short-form questionnaires of different lengths (1 to 8 items), resulting in 16 initial short-form questionnaires. These initial questionnaires were named after the algorithms and the number of items they contained. For example, AF-4 is the short-form questionnaire consisting of the first four items of the eight items identified using the ANOVA F-test. In comparison, SR-8 is the short-form questionnaire consisting of whole eight items of the eight items identified using Stepwise Regression.

Table 1. Ranking of feature importance using ANOVA F-tests.

Item number of BWAQ	F-value	p-value
29	144.42	0.000
7	139.02	0.000
21	121.43	0.000
9	113.42	0.000
12	111.52	0.000
5	107.33	0.000
33	107.04	0.000
14	101.99	0.000
32	95.03	0.000
31	89.48	0.000
22	85.48	0.000
23	78.79	0.000
11	76.98	0.000
17	74.72	0.000
13	67.59	0.000
30	65.80	0.000
16	65.04	0.000
20	61.14	0.000
8	57.78	0.000
1	57.73	0.000
4	55.79	0.000
10	55.43	0.000
34	53.03	0.000
6	53.00	0.000
15	47.13	0.000
2	46.40	0.000
3	45.77	0.000
18	43.27	0.000
25	43.24	0.000
27	40.16	0.000
24	36.40	0.000
28	19.15	0.000
26	6.53	0.011
19	4.53	0.034

The modeling results of the 16 initial short-form questionnaires by Logistic Regression, Support Vector Machine, Random Forest, and Naïve Bayes showed that the AUC of all models ranged from 0.72 to 0.98 (Table 2), it indicates that all the initial short-form questionnaires are effective in recognizing individual aggression. In general, the trained machine learning models perform better on the training set than the test set, and the smaller the difference in AUC between the models on the training set and the test set, the more stable their performance is and the closer their prediction results are to the actual situation. We further compare the differences in the AUC of different algorithms on training and test data (Figure 1) and find that the AUC of Random Forest is the most stable among all algorithms, which is more suitable for subsequent analysis.

Table 2. AUC of Machine Learning algorithms on initial version of short-form questionnaires.

Questionnaire	Logistics Regression		Support Vector Machine		Random Forest		Naïve Bayes	
	train	test	train	test	train	test	train	test
SR-1	0.83	0.72	0.83	0.72	0.83	0.72	0.83	0.72
SR-2	0.90	0.79	0.90	0.82	0.87	0.83	0.89	0.79
SR-3	0.93	0.83	0.92	0.83	0.87	0.81	0.92	0.83
SR-4	0.94	0.84	0.94	0.84	0.90	0.82	0.93	0.83
SR-5	0.95	0.86	0.95	0.87	0.92	0.87	0.94	0.85
SR-6	0.97	0.89	0.97	0.88	0.94	0.88	0.96	0.87
SR-7	0.98	0.91	0.98	0.91	0.95	0.87	0.97	0.89
SR-8	0.98	0.93	0.98	0.93	0.95	0.89	0.97	0.91
AF-1	0.83	0.72	0.83	0.72	0.83	0.72	0.83	0.72
AF-2	0.88	0.78	0.89	0.78	0.86	0.82	0.89	0.78
AF-3	0.93	0.81	0.92	0.80	0.85	0.79	0.93	0.80
AF-4	0.94	0.82	0.94	0.83	0.89	0.86	0.95	0.81
AF-5	0.95	0.85	0.95	0.85	0.90	0.88	0.95	0.85
AF-6	0.95	0.85	0.95	0.85	0.90	0.86	0.95	0.85
AF-7	0.95	0.87	0.96	0.87	0.93	0.91	0.95	0.86
AF-8	0.96	0.90	0.96	0.90	0.93	0.90	0.96	0.89

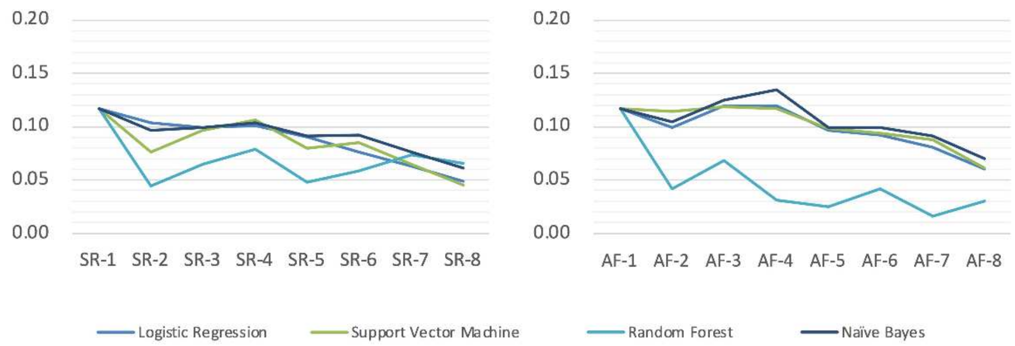

Figure 1. AUC differences between Machine Learning algorithms on training and test sets.

Among the 16 initial versions of short-form questionnaire models in Random Forest, there are three short-form questionnaires with AUC, accuracy, recall, and precision above 0.8, namely SR-5, AF-4, and AF-8 (Table 3). AF-4 has 50% fewer items than AF-8; AUC, accuracy, and precision have decreased by 4%, 1%, and 4%, respectively, and recall has increased by 3%. AF-4 has 20% fewer items than SR-5, 1% less AUC, and the same accuracy, recall, and precision. In contrast, AF-4 reduced the number of items by 20-50%, but the performance indicator changed by only 1-5%, and there was a high correlation with BWAQ (Pearson's correlation = 0.85, $p < 0.001$).

Table 3. Performance of the initial version of short-form questionnaires based on Random Forest.

Questionnaire	AUC	Accuracy	Recall	Precision
SR-1	0.72	0.77	0.85	0.77
SR-2	0.83	0.77	0.85	0.77
SR-3	0.81	0.72	0.80	0.74
SR-4	0.82	0.75	0.78	0.80
SR-5	0.87	0.81	0.83	0.85
SR-6	0.88	0.79	0.78	0.86
SR-7	0.87	0.81	0.78	0.89
SR-8	0.89	0.79	0.75	0.88
AF-1	0.72	0.77	0.85	0.77
AF-2	0.82	0.75	0.75	0.81
AF-3	0.79	0.75	0.78	0.80
AF-4	0.86	0.81	0.83	0.85
AF-5	0.88	0.77	0.75	0.83
AF-6	0.86	0.77	0.75	0.83
AF-7	0.91	0.81	0.78	0.89
AF-8	0.90	0.82	0.80	0.89

AF-4 includes four items, and its cutoff ranges from 4 to 20. We calculated the evaluation metrics corresponding to each cutoff according to the confuse matrix (Table 4). As the cutoff increases, AUC and accuracy first increase and then decrease, recall decreases from 1 to 0.04, and precision increases from 0.52 to 1 (Figure 2). When the cutoff of the AF-4 is changed, the relationship between recall and precision is in a trade-off relationship. Therefore, the performance of the questionnaire under a particular cutoff condition cannot be judged by either recall or precision alone; both recall and precision are essential and need to be considered. Therefore, to better estimate the cutoff, it is necessary to calculate the F1 score, which is the harmonic mean of recall and precision, to estimate the optimal cutoff for the AF-4 (Table 4). Comparing the AF-4 under 17 cutoff conditions, AF-4 has the best AUC, accuracy, and F1 score when the cutoff is 11. Therefore, the optimal short-form version of BWAQ based on the Machine Learning algorithm is AF-4 with a cutoff of 11, which we named BWAQ-ML.

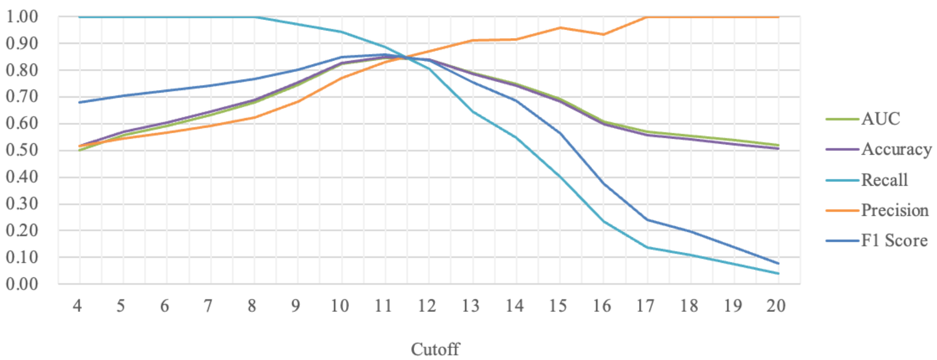

Figure 2. Trend of the evaluation metrics with cutoff.

Table 4. Evaluation metrics for different cutoffs of AF-4.

Cutoff	AUC	Accuracy	Recall	Precision	F1 Score
4	0.50	0.52	1.00	0.52	0.68
5	0.56	0.57	1.00	0.55	0.71
6	0.59	0.60	1.00	0.57	0.72
7	0.63	0.64	1.00	0.59	0.74

8	0.68	0.69	1.00	0.62	0.77
9	0.75	0.75	0.97	0.68	0.80
10	0.82	0.83	0.94	0.77	0.85
11	0.85	0.85	0.89	0.83	0.86
12	0.84	0.84	0.81	0.87	0.84
13	0.79	0.79	0.65	0.91	0.76
14	0.75	0.74	0.55	0.91	0.69
15	0.69	0.68	0.40	0.96	0.56
16	0.61	0.60	0.23	0.93	0.37
17	0.57	0.56	0.14	1.00	0.24
18	0.55	0.54	0.11	1.00	0.20
19	0.54	0.52	0.07	1.00	0.14
20	0.52	0.51	0.04	1.00	0.08

3.2. Validating BWAQ

Cronbach's alpha for the BWAQ-ML, RPQ, and PTM were 0.84, 0.83, and 0.93, respectively, and Pearson's correlation for the BWAQ-ML with the RPQ and PTM was 0.514 ($p < 0.001$) and -0.042 ($p = 0.375$), respectively, indicating that the BWAQ-ML can effectively measure individual aggression. We labeled samples scoring in the top 50% on the RPQ as high aggression and those scoring in the bottom 50% as low aggression. Comparing the classification results of the BWAQ-ML and the RPQ, the degree of overlap was 71%, indicating that the content measured by the BWAQ-ML was consistent with that of the RPQ. There was a significant positive correlation between the items of the BWAQ-ML, with correlation coefficients ranging from 0.469 to 0.621; there was a significant positive correlation between the items and the total score, with correlation coefficients ranging from 0.618 to 0.730 (Table 5), which indicates a good differentiation between the items.

Table 5. Correlation matrix for each item and total of BWAQ-ML.

	BWAQ-ML_1	**BWAQ-ML_2**	**BWAQ-ML_3**	**BWAQ-ML_4**	**BWAQ-ML_Total**
BWAQ-ML_1	1				
BWAQ-ML_2	.579**	1			
BWAQ-ML_3	.469**	.547**	1		
BWAQ-ML_4	.609**	.621**	.609**	1	
BWAQ-ML_Total	.618**	.702**	.677**	.730**	1

* Note: **. Correlation is significant at the 0.01 level (2-tailed).

4. Discussion

The number of items in the BWAQ-ML is only 12% of the BWAQ but retains 85% of the AUC and accuracy, 89% of the recall, and 83% of the precision of the full version of the questionnaire. In other words, the BWAQ-ML, with a reduced number of items by 88%, can identify 89% of the cases classified by the BWAQ as high aggression. The internal consistency coefficient of the short version of the questionnaire was 0.84, which was significantly positively correlated with the RPQ, and its categorization results reached a 71% overlap with the RPQ, with no significant correlation with the PTM, which indicated that the BWAQ-ML had good reliability and validity. Overall, the BWAQ-ML, which contains only four items, has sufficient validity as a condensed version of the BWAQ.

As the length of the questionnaire increases, participants experience IER due to carelessness or lack of effort [15], and the more backward the items, the more likely they are to experience IER [17], which can adversely affect the credibility of the study results [18]. The percentage of participants experiencing IER while completing the questionnaire can be up to 30% [16]. Therefore, the questionnaire should not be too long. The BWAQ-ML contains only four items, and the time required

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 August 2023 doi:10.20944/preprints202308.0189.v1

for participants to complete all the items is no more than 1 minute, which significantly reduces the perceived difficulty of the participants in completing the questionnaire, improves the completion rate and credibility of the questionnaire responses, and reduces the likelihood of the emergence of IER. In addition, by using the short-form questionnaire, the researcher could also measure more variables at once during the administration process. If necessary, the researcher can get the results quickly by oral calculation on the spot, which is easy and less prone to errors.

The BWAQ contains five aggressive behavioral manifestations: physical aggression, verbal aggression, hostility, anger, and indirect aggression [13]. The first two items of the BWAQ-ML belong to the anger dimension, and the last two belong to the hostility dimension. The positive correlation between hostility and suicide is more significant than the other dimensions [6,36]. Therefore, the short-form questionnaire may be more conducive to detecting highly aggressive individuals with suicidal tendencies. In addition, since these four items do not directly ask participants about the frequency of aggression, they reduce the social desirability of participants, which would facilitate the investigator's early identification of high-risk individuals for comprehensive intervention.

This study has the following limitations. First, the BWAQ-ML is simplified based on the BWAQ, so the performance of the short-form questionnaire relies on the original questionnaire. Second, the mean score of the BWAQ's norm for adolescents aged 12 to 18 ranged from 61 to 79, but the mean score for the simplification samples in this study was 90. There are two possible reasons for the higher mean score: 1) the urban norms were developed 12 years ago (2011), and people's scores on the BWAQ have changed; 2) this study was at the end of the secondary school year when the sample was administered, and exam pressure may have influenced subjects' scores on the BWAQ. Third, although Random Forest, which has a more stable AUC, is used as a reliable algorithm in the simplification phase, it is still possible that the Machine Learning algorithm may be overfitted due to the small sample size, which affects the results. Fourth, the BWAQ-ML has only four items, which is unsuitable for embedding recognition scales in the questionnaire for recognition of IER. In summary, it is recommended that future studies validate the BWAQ-ML using a larger sample.

5. Conclusions

The BWAQ-ML is a streamlined version of the BWAQ to measure adolescent aggression effectively. The items of BWAQ-ML are as follows:

(1) At times I feel like a bomb ready to explode.
(2) At times I get very angry for no good reason.
(3) I sometimes feel that people are laughing at me behind my back.
(4) I wonder why sometimes I feel so bitter about things.

Author Contributions: Conceptualization, J.L. and X.J.; methodology, X.J.; software, X.J.; validation, J.L., X.J. and Y.Y.; formal analysis, X.J.; investigation, Y.Y.; resources, Y.Y.; data curation, X.J.; writing—original draft preparation, X.J.; writing—review and editing, J.L.; visualization, X.J.; supervision, J.L.; project administration, J.L.

Funding: This research was funded by the Institute of Psychology, CAS, grant number GJ202011, Chengdu Office of Philosophy and Social Science, grant number 2022CZ083, and Sichuan Provincial Center for Educational Informationization and Big Data, grant number DSJ2022050.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Ethics Committee of Sichuan Normal University (2023LS029).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.
Data Availability Statement: The data presented in this study are openly available in OSF Registries at https://osf.io/znr46/.

Acknowledgments: The authors thank all the volunteers, teachers, and administrative supporters who help to conduct the investigations.

Conflicts of Interest: The authors declare no conflict of interest.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 August 2023 doi:10.20944/preprints202308.0189.v1

References

1. Ma, Y.; Jiao, T.; Guo, S.; Tang, J. Association of aggression and social support among adolescents. *Chinese Journal of School Health* **2022**, *43*, 671-675, doi:10.16835/j.cnki.1000-9817.2022.05.008.
2. Carlyle, K.E.; Steinman, K.J. Demographic Differences in the Prevalence, Co-Occurrence, and Correlates of Adolescent Bullying at School. *Journal of School Health* **2007**, *77*, 623-629, doi:https://doi.org/10.1111/j.1746-1561.2007.00242.x.
3. Wu, P.; Liu, H.; Chen, J.; Xie, J. The Transition of Aggressive Patterns among Youth: An Application of Latent Transition Analysis. *Journal of Psychdogical Science* **2014**, *37*, 1167-1173, doi:10.16719/j.cnki.1671-6981.2014.05.023.
4. Schlomer, G.L.; Cleveland, H.H.; Vandenbergh, D.J.; Feinberg, M.E.; Neiderhiser, J.M.; Greenberg, M.T.; Spoth, R.; Redmond, C. Developmental Differences in Early Adolescent Aggression: A Gene × Environment × Intervention Analysis. *Journal of Youth and Adolescence* **2015**, *44*, 581-597, doi:10.1007/s10964-014-0198-4.
5. Tang, J.; Ma, Y.; Guo, Y.; Ahmed, N.I.; Yu, Y.; Wang, J. Association of Aggression and Non-Suicidal Self Injury: A School-Based Sample of Adolescents. *PLOS ONE* **2013**, *8*, e78149, doi:10.1371/journal.pone.0078149.
6. Peng, C.; Guo, T.; Cheng, J.; Wang, M.; Tan, Y.; Rong, F.; Kang, C.; Ding, H.; Wang, Y.; Yu, Y. Association between childhood physical abuse and suicidal behaviors among Chinese adolescents: The mediation of aggression. *Journal of Affective Disorders* **2022**, *318*, 338-346, doi:https://doi.org/10.1016/j.jad.2022.09.021.
7. Zhang, Y.; Wu, C.; Yuan, S.; Xiang, J.; Hao, W.; Yu, Y. Association of aggression and suicide behaviors: A school-based sample of rural Chinese adolescents. *Journal of Affective Disorders* **2018**, *239*, 295-302, doi:https://doi.org/10.1016/j.jad.2018.07.029.
8. Preddy, T.M.; Fite, P.J. Differential Associations Between Relational and Overt Aggression and Children's Psychosocial Adjustment. *Journal of Psychopathology and Behavioral Assessment* **2012**, *34*, 182-190, doi:10.1007/s10862-011-9274-1.
9. Vuoksimaa, E.; Rose, R.J.; Pulkkinen, L.; Palviainen, T.; Rimfeld, K.; Lundström, S.; Bartels, M.; van Beijsterveldt, C.; Hendriks, A.; de Zeeuw, E.L.; et al. Higher aggression is related to poorer academic performance in compulsory education. *Journal of Child Psychology and Psychiatry* **2021**, *62*, 327-338, doi:https://doi.org/10.1111/jcpp.13273.
10. Fite, P.J.; Hendrickson, M.; Rubens, S.L.; Gabrielli, J.; Evans, S. The Role of Peer Rejection in the Link between Reactive Aggression and Academic Performance. *Child & Youth Care Forum* **2013**, *42*, 193-205, doi:10.1007/s10566-013-9199-9.
11. Muñoz Reyes, J.A.; Guerra, R.; Polo, P.; Cavieres, E.; Pita, M.; Turiégano, E. Using an evolutionary perspective to understand the relationship between physical aggression and academic performance in late adolescents. *Journal of School Violence* **2019**, *18*, 39-48, doi:10.1080/15388220.2017.1368397.
12. Buss, A.H.; Perry, M. The Aggression Questionnaire. *Journal of Personality and Social Psychology* **1992**, *63*, 452-459, doi:10.1037/0022-3514.63.3.452.
13. Buss, A.H.; Warren, W.L. *The aggression questionnaire manual*; Western Psychological Services: Los Angeles, 2000.
14. Huang, J.; Tang, J.; Tang, L.; Chang, H.J.; Ma, Y.; Yan, Q.; Yu, Y. Aggression and related stressful life events among Chinese adolescents living in rural areas: A cross-sectional study. *Journal of Affective Disorders* **2017**, *211*, 20-26, doi:https://doi.org/10.1016/j.jad.2016.12.044.
15. Curran, P.G. Methods for the detection of carelessly invalid responses in survey data. *Journal of Experimental Social Psychology* **2016**, *66*, 4-19, doi:https://doi.org/10.1016/j.jesp.2015.07.006.
16. Zhong, X.; Li, M.; Li, L. Preventing and detecting insufficient effort survey responding. *Advances in Psychological Science* **2021**, *29*, 225-237, doi:10.3724/sp.J.1042.2021.00225.
17. Galesic, M.; Bosnjak, M. Effects of Questionnaire Length on Participation and Indicators of Response Quality in a Web Survey. *Public Opinion Quarterly* **2009**, *73*, 349-360, doi:10.1093/poq/nfp031.
18. Maniaci, M.R.; Rogge, R.D. Caring about carelessness: Participant inattention and its effects on research. *Journal of Research in Personality* **2014**, *48*, 61-83, doi:https://doi.org/10.1016/j.jrp.2013.09.008.
19. Maxwell, J.P. Development and Preliminary Validation of a Chinese Version of the Buss–Perry Aggression Questionnaire in a Population of Hong Kong Chinese. *Journal of Personality Assessment* **2007**, *88*, 284-294, doi:10.1080/00223890701317004.
20. Kiewitz, C.; Weaver Iii, J.B. The aggression questionnaire. In *Handbook of research on electronic surveys and measurements.*; Idea Group Reference/IGI Global: Hershey, PA, US, 2007; pp. 343-347.
21. Lin, G.-H.; Li, C.-Y.; Sheu, C.-F.; Huang, C.-Y.; Lee, S.-C.; Huang, Y.-H.; Hsieh, C.-L. Using Machine Learning to Develop a Short-Form Measure Assessing 5 Functions in Patients With Stroke. *Archives of Physical Medicine and Rehabilitation* **2022**, *103*, 1574-1581, doi:https://doi.org/10.1016/j.apmr.2021.12.006.
22. Wang, I.; Li, P.-C.; Lee, S.-C.; Lee, Y.-C.; Wang, C.-H.; Hsieh, C.-L. Development of a Berg Balance Scale Short-Form Using a Machine Learning Approach in Patients With Stroke. *Journal of Neurologic Physical Therapy* **2023**, *47*, 44-51, doi:10.1097/npt.0000000000000417.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 August 2023 doi:10.20944/preprints202308.0189.v1

23. Lee, W.; Kim, H.; Shim, J.; Kim, D.; Hyeon, J.; Joo, E.; Joo, B.-E.; Oh, J. The simplification of the insomnia severity index and epworth sleepiness scale using machine learning models. *Scientific Reports* **2023**, *13*, 6214, doi:10.1038/s41598-023-33474-8.
24. Orrù, G.; De Marchi, B.; Sartori, G.; Gemignani, A.; Scarpazza, C.; Monaro, M.; Mazza, C.; Roma, P. Machine learning item selection for short scale construction: A proof-of-concept using the SIMS. *The Clinical Neuropsychologist* **2022**, 1-18, doi:10.1080/13854046.2022.2114548.
25. Morrison, A.S.; Ustun, B.; Horenstein, A.; Kaplan, S.C.; de Oliveira, I.R.; Batmaz, S.; Gross, J.J.; Sadikova, E.; Hemanny, C.; Pires, P.P.; et al. Optimized short-forms of the Cognitive Distortions Questionnaire. *Journal of Anxiety Disorders* **2022**, *92*, 102624, doi:https://doi.org/10.1016/j.janxdis.2022.102624.
26. Lin, G.-H.; Huang, C.-Y.; Lee, S.-C.; Chen, K.-L.; Lien, J.-J.J.; Chen, M.-H.; Huang, Y.-H.; Hsieh, C.-L. A 10-item Fugl-Meyer Motor Scale Based on Machine Learning. *Physical Therapy* **2021**, *101*, doi:10.1093/ptj/pzab036.
27. Maxwell, J.P. Psychometric properties of a Chinese version of the Buss–Warren Aggression Questionnaire. *Personality and Individual Differences* **2008**, *44*, 943-953, doi:https://doi.org/10.1016/j.paid.2007.10.037.
28. Zhang, P.; Yu, Y.; Liu, Z.; Meng, X. Chinese Norm of Buss-Warren Aggression Questionnaire from a standardized school-based sample aged 9 - 18 in urban areas. *School of Public Health* **2011**, *32*, 897-900, doi:10.16835/j.cnki.1000-9817.2011.08.002.
29. Wang, M.; Peng, C.; Chang, H.; Yu, M.; Rong, F.; Yu, Y. Interaction between Sirtuin 1 (SIRT1) polymorphisms and childhood maltreatment on aggression risk in Chinese male adolescents. *Journal of Affective Disorders* **2022**, *309*, 37-44, doi:https://doi.org/10.1016/j.jad.2022.04.063.
30. Raine, A.; Dodge, K.; Loeber, R.; Gatzke-Kopp, L.; Lynam, D.; Reynolds, C.; Stouthamer-Loeber, M.; Liu, J. The reactive–proactive aggression questionnaire: differential correlates of reactive and proactive aggression in adolescent boys. *Aggressive Behavior* **2006**, *32*, 159-171, doi:https://doi.org/10.1002/ab.20115.
31. Quan, F.; Xia, L. The Prediction of Hostile Attribution Bias on Reactive Aggression and the Mediating Role of Revenge Motivation. *Journal of Psychological Science* **2019**, *42*, 1434-1440, doi:10.16719/j.cnki.1671-6981.20190623.
32. Wang, Y.; Xiong, Y.; Ren, P.; Yang, L.; Miao, W. Effects of Bully-victimization on Proactive and Reactive Aggression in Early Adolescence: The Role of Moral Disengagement and Gender. *Psychological Development and Education* **2023**, *39*, 410-418, doi:10.16187/j.cnki.issn1001-4918.2023.03.12.
33. Li, R.; Xia, L. The mediating effect of aggression motivation on the relationship between trait anger and reactive aggression: A longitudinal study. *Acta Psychologica Sinica* **2021**, *53*, 788-797, doi:10.3724/SP.J.1041.2021.00788.
34. Carlo, G.; Randall, B.A. The Development of a Measure of Prosocial Behaviors for Late Adolescents. *Journal of Youth and Adolescence* **2002**, *31*, 31-44, doi:10.1023/A:1014033032440.
35. Yin, K.; Wu, X.; Zhang, T. Relationship between psychological privilege and prosocial behavior in college students in the post-epidemic period : A moderated chain mediation model. *China Journal of Health Psychology* **2022**, *30*, 1549-1554, doi:10.13342/j.cnki.cjhp.2022.10.022.
36. Zhang, P.; Roberts, R.E.; Liu, Z.; Meng, X.; Tang, J.; Sun, L.; Yu, Y. Hostility, Physical Aggression and Trait Anger as Predictors for Suicidal Behavior in Chinese Adolescents: A School-Based Study. *PLOS ONE* **2012**, *7*, e31044, doi:10.1371/journal.pone.0031044.