International Seminar on Student Research in Education, Science, and Technology
ISSN 3047-1532| Volume 2 April 2025, pp. 126-133

https://www.issrestec.id/
http://journal.ummat.ac.id/index.php/issrestec
Universitas Muhammadiyah Mataram, Indonesia

From Bystander to Defender: Developing and Validating a Scale for Measuring Defending Behaviors Among Indonesian Adolescents in Bullying Contexts

Khairatul Ulya Ilma[1], Tina Hayati Dahlan[2], Nandang Budiman[3]
[1]Educational Psychology, Universitas Pendidikan Indonesia, Indonesia, khairatul2ya.ilma@upi.edu
[2]Educational Psychology, Universitas Pendidikan Indonesia, Indonesia, tinadahlan_psi@upi.edu
[3]Educational Psychology, Universitas Pendidikan Indonesia, Indonesia, nandang.budiman@upi.edu

Abstract: Defending behaviors play a crucial role in mitigating the negative effects of bullying on victims. However, many prevention and intervention programs promote peer defending without a clear understanding of the specific behaviors adolescent's exhibit. To address this gap, a scale on defending behavior in bullying contexts among Indonesian adolescents was developed and validated using the Rasch Model. Rasch analysis, conducted with Winsteps software version 3.73, examined key psychometric properties, including discrimination index, including discrimination index, item difficulty, item-person fit, scale reliability, response category effectiveness, item bias, and respondent classification profiling. Data were collected from 142 junior high school students in Grades 7-9. The final version of the scale comprises 28 items distributed across four dimensions: solution focused defending, aggressive defending, comforting, and reporting to authority with 5 response options. The overall person-item interaction yielded a Cronbach's alpha of 0.88, indicating excellent internal consistency. Additionally, the person reliability index was 0.86, reflecting high response consistency among participants. These findings confirm that the scale meets psychometric requirements and is a valid and reliable tool for assessing defending behavior in adolescent bullying contexts.

Keywords: Defending Behavior, Bullying, Peer Intervention, Adolescents, Validity, Rasch Model.

Article History:
Received: 01-04-2025
Online : 24-04-2025

A. INTRODUCTION

Defending is conceptualized as a multidimensional construct that encompasses both direct and indirect forms of prosocial behavior aimed at supporting individuals who are being victimized (Lambe & Craig, 2020). These behaviors encompass a range of prosocial actions intended to stop bullying, such as directly intervening, seeking help from adults, or providing emotional support to the victim (Salmivalli, 2010). Such behaviors not only have the potential to interrupt bullying incidents but also contribute to the overall reduction of bullying within school settings. Moreover, when individuals who witness bullying, commonly referred to as bystanders, engage in defending, they may offer social protection against aggression from perpetrators, which in turn may reduce bullying over time (McDougall & Vaillancourt, 2015).

Bystanders who support victims or report bullying incidents to adults can enhance victims' psychological well-being (Polanin et al., 2012). Although defending does not always succeed in stopping bullying, defenders who offer support tend to generate positive outcomes, such as reduced anxiety and depression among victims (Salmivalli, 2014), and an increased sense of school connectedness (Laninga-Wijnen et al., 2022). As a result, many many prevention and intervention programs on decreasing peer victimization by promoting greater engagement in defending behaviors among students (Polanin et al., 2012). However, the

effectiveness of these efforts may be constrained without a comprehensive understanding of what defending behavior entails or how various defending strategies may differentially impact youth (Lambe & Craig, 2020).

In light of these considerations, the present study sought to investigate the underlying structure and dimensions of defending behaviors in the context of adolescent bullying, with the aim of advancing a more nuanced and comprehensive understanding of these behaviors within bullying research. In line with this objective, a psychometrically sound instrument was developed to measure defending behavior. The scale's reliability and validity were subsequently examined in a sample of Indonesian adolescents.

B. METHOD

This study was conducted in Indonesia in 2025 and involved a total of 142 junior high school students, comprising 102 girls and 40 boys, aged between 11 and 16 years. The demographic characteristics of the participants are presented in the table below.

Table 1. Demographic Profile of Participants

	Category	Frequency	Percentage
Gender	Boy	40	28.17
	Girl	102	71.83
Age	11	1	0.70
	12	20	14.08
	13	51	35.92
	14	58	40.85
	15	11	7.75
	16	1	0.70

The 32-item questionnaire was designed to investigate the dimensions of defending behaviors, conceptualized as a multidimensional construct encompassing both direct and indirect forms of defending. The development of the questionnaire was grounded in relevant theoretical frameworks, empirical research, and existing literature, with particular reference to the instrument developed by Lambe & Craig (2020). Generally, items were classified into 4 different sub-dimensions namely: 1. Solution-focused defending; 2. Aggressive defending; 3. Comforting; and 4. Reporting to authority (see table 2).

Table 2. Defending Subdimensions

Dimension	Subdimension	Description	Representative Item
Direct defending	Solution-focused defending	Students direct their attention to the bullying incident and actively seek solutions to address the situation experienced by the victim, with the aim of preventing the bullying from escalating further.	I encouraged the bully to apologize to the person they bullied.

	Aggressive defending	Aggressive defending occurs when a student responds to bullying by confronting the bully with threats or aggressive behavior in an attempt to force them to stop or reverse their harmful actions.	I stepped in and pushed the bully away to protect the peer being bullied.
Indirect defending	Comforting	Comforting refers to to the emotional support and reassurance offered by individuals to victims of bullying, typically aimed at helping them feel seen, safe, and less alone following the incident.	I supported the victim of bullying by offering emotional comfort, such as providing a hug or offering words of encouragement to help them feel more at ease.
	Reporting to authority	The student who witnesses bullying reports the incident to an adult, such as a teacher, who is able to intervene and address the situation.	I immediately report the bullying to the teacher.

Each item was thoughtfully designed to encompass a wide range of prosocial strategies that adolescents may employ when defending peers who are targets of bullying. Responses were collected using a five-point Likert scale, ranging from strongly inappropriate to strongly appropriate. The data analysis was conducted using the Rasch model, applied through the Winstep software.

C. RESULTS AND DISCUSSION

1. Instrument analysis

Based on the data presented in the table below, the instrument demonstrated a test reliability coefficient of 0.88, this indicates that the overall interaction between persons and items falls within the "excellent" category. With a person reliability of 0.86 and an item reliability of 0.97, it can be concluded that the consistency of respondents' answers is strong, and the quality of the items included in the instrument is classified as outstanding. According to Boone et al. (2014), person reliability reflects how well a test is able to distinguish individuals into specific categories. A reliability coefficient of 0.8 suggests that the test can differentiate individuals into approximately two or three distinct categories. Moreover, a high item reliability value also indicates that the sample used in the test is capable of providing accurate estimates of item positioning or item difficulty levels on the latent variable being measured.

Tabel 3. Instrument Analysis

	Means	**SD**	**Separation**	**Reliability**	**Cronbach Alpha**
Person	0.05	0.51	2.46	0.86	0.88
Item	0.08	0.44	5.31	0.97	

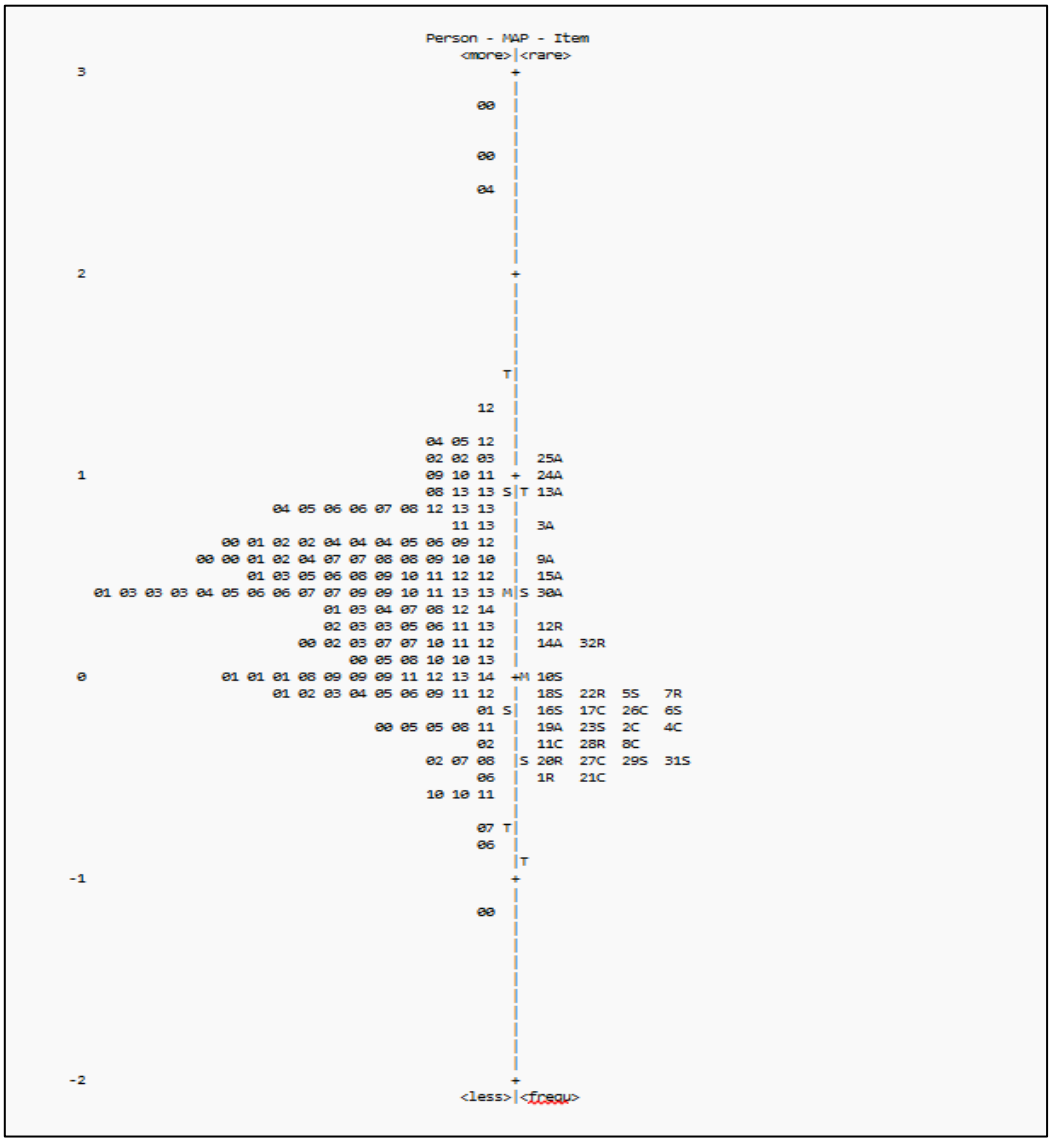

Figure 1. Person and Item Map

The average item logit value is always set at 0.0 logit, which serves as the initial reference point of the scale (Sumintono & Widhiarso, 2015). On the Wright Map, when the average person logit is higher than the average item difficulty, it indicates that most respondents found the items in the instrument relatively easy to agree with or respond to (Boone et al., 2014). Figure 1 shows that the average person logit is +0.38, which is slightly above the item average of 0.0 logit. This position suggests that respondents tend to agree with or display the measured characteristic (defending behavior) at a slightly higher level than the average difficulty (i.e., the level of endorsement or effort required to respond positively to the items). In other words, the defending behavior instrument being developed appears to be more accessible for most respondents.

Moreover, the item map reveals that a large number of items are concentrated below the item mean logit value (0.0), indicating that many items share a similar (easy) level of difficulty. There are also noticeable gaps along the higher and lower ability ranges, suggesting underrepresentation at those levels. The item map in Figure 1 further shows that item No. 25A (+1.06 logit) is the most difficult item to endorse, while items No. 1R and 28C (-0.50 & -0.48 logit) are the easiest for respondents to endorse.

2. Item Conformity Level

Item fit refers to whether a questionnaire item functions appropriately in measuring the intended construct. An item that does not fit is indicative of potential misconceptions or misunderstandings by respondents regarding the statement (Sumintono & Widhiarso, 2015). o identify which items are fitting or misfitting, the INFIT MNSQ value of each item is examined. The average value and standard deviation are summed, and items with logit values exceeding this threshold are considered misfitting. Another criterion for assessing item fit is provided by Boone et al. (2014), hich includes the following acceptable ranges: MNSQ values between 0.5 and 1.5, ZSTD values between -2.0 and +2.0, and Point Measure Correlation values between 0.4 and 0.85. When an item does not meet the thresholds for MNSQ and Point Measure Correlation but still falls within the acceptable range for ZSTD, the item can still be considered fitting. This indicates that the item may be retained in the instrument without requiring any modifications. The results of the Rasch Model analysis for the defending behavior instrument are illustrated in the figure 2 below.

Item STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.		EXACT MATCH OBS%	EXACT MATCH EXP%	Item
25	304	141	1.06	.08	1.44	3.6	2.07	6.5	A	.14	.48	33.3	31.0	25A
24	312	141	1.01	.08	1.54	4.4	1.81	5.4	B	.21	.49	34.0	30.0	24A
3	353	141	.78	.07	1.50	4.4	1.71	5.4	C	.11	.49	22.7	29.3	3A
13	335	141	.88	.07	1.22	2.1	1.68	4.9	D	.24	.49	30.5	28.7	13A
1	576	141	-.50	.09	1.39	2.7	1.25	1.6	E	.47	.40	30.5	37.2	1R
8	547	140	-.32	.08	1.22	1.7	1.32	2.2	F	.36	.42	32.9	34.9	8C
15	410	141	.48	.07	1.24	2.3	1.30	2.6	G	.19	.48	30.5	29.3	15A
18	510	141	-.06	.08	.71	-2.9	1.24	1.9	H	.48	.45	48.9	32.4	18S
9	382	141	.62	.07	1.23	2.3	1.22	1.9	I	.33	.49	32.6	28.9	9A
14	466	141	.19	.07	1.18	1.7	1.15	1.3	J	.45	.47	27.0	30.9	14A
32	470	141	.17	.07	1.13	1.2	1.12	1.0	K	.45	.47	35.5	31.2	32R
2	545	141	-.28	.08	1.05	.4	1.00	.1	L	.49	.43	34.0	34.7	2C
12	452	141	.26	.07	.92	-.8	1.03	.3	M	.44	.47	34.0	30.4	12R
5	519	141	-.11	.08	.99	.0	1.03	.3	N	.44	.45	32.6	32.7	5S
30	417	141	.44	.07	1.00	.0	.96	-.3	O	.45	.48	35.5	29.4	30A
6	531	141	-.18	.08	.99	.0	.97	-.2	P	.54	.44	27.7	33.2	6S
4	546	141	-.28	.08	.97	-.2	.98	-.1	p	.45	.43	38.3	34.7	4C
23	546	141	-.28	.08	.94	-.5	.90	-.7	o	.49	.43	36.2	34.7	23S
10	499	141	.01	.08	.93	-.6	.93	-.5	n	.44	.46	42.6	32.0	10S
11	558	141	-.37	.08	.92	-.6	.91	-.6	m	.47	.42	36.9	35.8	11C
20	560	141	-.38	.09	.85	-1.2	.91	-.6	l	.54	.42	42.6	35.8	20R
29	565	141	-.42	.09	.89	-.8	.82	-1.3	k	.62	.41	36.9	35.9	29S
16	529	141	-.17	.08	.89	-1.0	.84	-1.2	j	.52	.44	41.1	33.1	16S
28	559	141	-.37	.08	.88	-.9	.81	-1.4	i	.60	.42	33.3	35.8	28R
17	532	141	-.19	.08	.86	-1.2	.78	-1.8	h	.60	.44	44.0	33.4	17C
22	520	141	-.12	.08	.83	-1.5	.85	-1.2	g	.53	.44	40.4	32.7	22R
26	524	141	-.14	.08	.80	-1.9	.84	-1.3	f	.50	.44	43.3	32.8	26C
7	517	141	-.10	.08	.78	-2.1	.79	-1.8	e	.57	.45	43.3	32.7	7R
31	565	141	-.42	.09	.78	-1.9	.76	-1.8	d	.58	.41	41.1	35.9	31S
21	573	141	-.48	.09	.70	-2.5	.68	-2.4	c	.55	.40	45.4	37.1	21C
27	567	141	-.43	.09	.66	-3.0	.65	-2.8	b	.54	.41	49.6	36.6	27C
19	547	141	-.29	.08	.60	-3.8	.58	-3.6	a	.63	.43	44.0	34.7	19A
MEAN	494.9	141.0	.00	.08	1.00	.0	1.06	.4				36.9	33.1	
S.D.	80.0	.2	.45	.01	.24	2.1	.34	2.4				6.4	2.5	

Figure 2. Item fit order

Based on the analysis results, the sum of the item logit mean and standard deviation in the table is 1.00 + 0.24 = +1.24. According to this criterion, four items have INFIT MNSQ values exceeding this threshold: item 25A (+1.44), 24A (+1.54), 3A (+1.50), and 1R (+1.39). Furthermore, when considering the values of OUTFIT MNSQ, OUTFIT ZSTD, and Point Measure Correlation, items 25A, 24A, 3A, and 13A tend to be misfitting, as they do not meet the requirements for all three indicators. Therefore, these items require further evaluation to determine whether they should be removed or retained with modifications. Meanwhile, items 8C, 15A, and 9A can be retained without revision. Items 8C and 15A have OUTFIT MNSQ values that remain within acceptable tolerance limits, while item 9A only fails to meet the Point Measure Correlation criterion, thus still qualifying as a fitting item.

3. Rating Scale Diagnostic

This diagnostic analysis is conducted to determine whether participants are able to distinguish between the response categories on a 5-point scale (1, 2, 3, 4, and 5). When the observed averages follow an ascending order across response categories, it indicates that respondents can meaningfully differentiate among the options. The Andrich threshold values, which reflect the transition points between adjacent categories, are presented in Figure 3.

SUMMARY OF CATEGORY STRUCTURE. Model="R"

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	ANDRICH THRESHOLD	CATEGORY MEASURE	
1	1	549	12	-.25	-.41	1.23	1.26	NONE	(-1.83)	1
2	2	496	11	-.20	-.07	.76	.80	-.14	-.73	2
3	3	909	20	.14	.24	.86	.82	-.52	-.06	3
4	4	1217	27	.54	.52	.82	.86	.09	.68	4
5	5	1340	30	.87	.84	1.04	1.19	.57	(2.00)	5
MISSING		1	0	1.45						

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

Figure 3. Rating Scale Diagnostic

Figure 3 indicates that the response categories 1, 2, 4, and 5 demonstrate a consistent and orderly progression, suggesting that participants were able to distinguish among these options effectively. However, response category 3 appears to cause confusion among respondents, as evidenced by irregularities in its threshold values. This inconsistency in the functioning of category 3 disrupts the overall uniformity of the rating scale. Consequently, while participants seem to clearly differentiate between response options 1, 2, 4, and 5, category 3 is less clearly understood and may not function as intended.

This study developed a scale to measure defending behaviors among adolescents in the context of bullying. The scale demonstrated satisfactory psychometric properties, confirming its reliability and validity as an instrument for assessing adolescents' defending behaviors in bullying situations. The scale consists of four subdimensions: solution-focused defending, aggressive defending, comforting, and reporting to authority. The results of the reliability analysis indicated that the defending behavior scale exhibited high internal consistency.

According to the Wright Map, the most challenging item for participants to endorse was item number 25, which corresponds to a statement in the *aggressive defending* dimension, a direct form of defending behavior. This finding aligns with previous research indicating that aggressive defending is often perceived as a less favorable or less socially acceptable strategy among adolescents, as it involves confrontation and potential escalation of conflict (Pozzoli & Gini, 2010; Thornberg et al., 2012). Adolescents may hesitate to engage in or endorse aggressive defending due to concerns about retaliation, peer disapproval, or violation of school norms that discourage aggressive responses—even if such responses are intended to protect the victim.

In contrast, the items most easily endorsed by participants were items 1 and 28, which fall under the reporting to authority dimension—an indirect form of defending behavior. This finding is consistent with prior studies indicating that reporting to authority figures, such as teachers or school staff, is among the most commonly used strategies when adolescents

witness bullying (Casey et al., 2017; Chen et al., 2024). Students often perceive this approach as low-risk and socially acceptable(Derr & Morrow, 2020; Pozzoli & Gini, 2010). Moreover, Thornberg et al. (2012) hat students believe adults possess both the power and responsibility to stop bullying, reinforcing their preference for reporting over direct intervention. Wachs et al. (2019), further emphasized that students' perceptions of the effectiveness of adult intervention, along with previous positive experiences, significantly influence their willingness to report bullying incidents.

In connection with these findings, the analysis of response patterns also revealed that response options 1, 2, 4, and 5 were generally acceptable to participants; however, response option 3 appeared less comprehensible. The difficulty in interpreting the midpoint response may be attributed to cultural influences, particularly in high-context cultures such as Indonesia (Dolnicar & Grün, 2011). n such settings, students are often less accustomed to responding to rating scales with numerous alternatives, as the cognitive-emotional interplay in decision-making may interfere with their ability to distinguish subtle gradations in response options (Ilfiandra et al., 2022).

D. CONCLUSIONS AND SUGGESTIONS

Based on the Rasch Model analysis, the instrument demonstrated a high level of reliability, with a test reliability score of 0.88. Referring to Table 1, the interaction between persons and items can be categorized as excellent. The person reliability value of 0.86 and item reliability value of 0.97 indicate that respondents provided consistent responses, and that the quality of the items is classified as outstanding. However, the item fit analysis revealed that items 25A, 24A, 3A, and 13A were identified as misfitting, as they did not meet the three criteria of OUTFIT MNSQ, OUTFIT ZSTD, and Point Measure Correlation. These items require further review to determine whether they should be revised or removed to enhance the overall quality of the instrument. In addition, since response option 3 was not well understood by respondents, it is recommended to limit the response scale to four clearly distinguishable answer choices out of the original five-point scale.

Although the defending behavior scale has met the psychometric property requirements and can be utilized as a tool to measure students' defending responses in bullying situations, several limitations remain. One notable limitation lies in the relatively small research sample size, which may affect the stability and generalizability of the findings. Future research is advised to employ a larger participant-to-item ratio in order to enhance data stability and strengthen the reliability of the scale. Moreover, a single cross-sectional study is insufficient to comprehensively develop and validate an instrument that consistently produces strong psychometric results. As such, future research is encouraged to adopt longitudinal designs and incorporate qualitative data to better capture the developmental dynamics of defending behavior across different contexts.

It is also important to acknowledge that this scale employs a self-report format, which primarily captures students' perceptions of their defending behaviors. However, it lacks the capacity to fully explain the underlying psychological processes that influence each individual's responses. Therefore, further research is recommended to explore the protective factors that contribute to defending behavior, allowing for a clearer understanding of its configuration and determinants. Additionally, the context within each item statement remains relatively general. Future investigations should consider incorporating specific situational contexts into the items to reflect more accurate perceptions of students' defending behavior, thereby enriching the depth and variation of defending behavior research.

ACKNOWLEDGMENTS

We would like to express our sincere gratitude to Universitas Pendidikan Indonesia for providing the research grant that supported this study. Our appreciation also extends to all individuals who contributed to the completion of this research.

REFERENCES

Boone, W. J., Yale, M. S., & Staver, J. R. (2014). Rasch analysis in the human sciences. In *Rasch Analysis in the Human Sciences*. https://doi.org/10.1007/978-94-007-6857-4

Casey, E. A., Storer, H. L., Herrenkohl, T. I., Casey, E. A., & Storer, H. L. (2017). *Mapping a Continuum of Adolescent Helping and*.

Chen, H., Chen, C., Li, Y., & Fan, C. (2024). *behavioral sciences Development and Validation of the Defending Behavior Scale of Cyberbullying for Adolescents*.

Derr, S., & Morrow, M. T. (2020). Effects of a Growth Mindset of Personality on Emerging Adults' Defender Self-Efficacy, Moral Disengagement, and Perceived Peer Defending. *Journal of Interpersonal Violence*, *35*(3–4), 542–570. https://doi.org/10.1177/0886260517713716

Dolnicar, S., & Grün, B. (2011). *Three good reasons NOT to use five and seven point Likert items What makes Airbnb so different? View project Bibliotherapy-Cognitive behavioural therapy in treatment-resistant depression View project*. http://ro.uow.edu.au/commpapers/775

Ilfiandra, Nadhirah, N. A., Suryana, D., & binti Ahmad, A. (2022). Development and Validation Peaceful Classroom Scale: Rasch Model Analysis. *International Journal of Instruction*, *15*(4), 497–514. https://doi.org/10.29333/iji.2022.15427a

Lambe, L. J., & Craig, W. M. (2020). Peer defending as a multidimensional behavior: Development and validation of the Defending Behaviors Scale. *Journal of School Psychology*, *78*(December 2018), 38–53. https://doi.org/10.1016/j.jsp.2019.12.001

Laninga-Wijnen, L., van den Berg, Y. H. M., Garandeau, C. F., Mulder, S., & de Castro, B. O. (2022). Does Being Defended Relate to Decreases in Victimization and Improved Psychosocial Adjustment Among Victims? *Journal of Educational Psychology*, *115*(2), 363–377. https://doi.org/10.1037/edu0000712

Matell, M. S., & Jacoby, J. (1971). Is there an optimal number of alternatives for likert scale items? study 1: Reliability and validity. *Educational and Psychological Measurement*, *31*(3), 657–674. https://doi.org/10.1177/001316447103100307

McDougall, P., & Vaillancourt, T. (2015). Long-term adult outcomes of peer victimization in childhood and adolescence: Pathways to adjustment and maladjustment. *American Psychologist*, *70*(4), 300–310. https://doi.org/10.1037/a0039174

Polanin, J. R., Espelage, D. L., & Pigott, T. D. (2012). A meta-analysis of school-based bullying prevention programs' effects on bystander intervention behavior. *School Psychology Review*, *41*(1), 47–65. https://doi.org/10.1080/02796015.2012.12087375

Pozzoli, T., & Gini, G. (2010). Active defending and passive bystanding behavior in bullying: The role of personal characteristics and perceived peer pressure. *Journal of Abnormal Child Psychology*, *38*(6), 815–827. https://doi.org/10.1007/s10802-010-9399-9

Salmivalli, C. (2010). Bullying and the peer group: A review. *Aggression and Violent Behavior*, *15*(2), 112–120. https://doi.org/10.1016/j.avb.2009.08.007

Salmivalli, C. (2014). Participant Roles in Bullying: How Can Peer Bystanders Be Utilized in Interventions? *Theory into Practice*, *53*(4), 286–292. https://doi.org/10.1080/00405841.2014.947222

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Model Rasch untuk Penelitian Ilmu-ilmu Sosial.* Penerbit Trim Komunika.

Thornberg, R., Tenenbaum, L., Varjas, K., Meyers, J., Jungert, T., & Vanegas, G. (2012). Bystander motivation in bullying incidents: To intervene or not to intervene? *Western Journal of Emergency Medicine*, *13*(3), 247–252. https://doi.org/10.5811/westjem.2012.3.11792

Wachs, S., Bilz, L., Niproschke, S., & Schubarth, W. (2019). Bullying Intervention in Schools: A Multilevel Analysis of Teachers' Success in Handling Bullying From the Students' Perspective. *Journal of Early Adolescence*, *39*(5), 642–668. https://doi.org/10.1177/0272431618780423