Principal Component Analysis and Exploratory Factor Analysis of the Mechanical Waves Conceptual Survey

Purwoko Haryadi Santoso[1,2], Edi Istiyono[1,3], Haryanto[1]

Graduate School of Educational Research and Evaluation, Universitas Negeri Yogyakarta, Indonesia[1]
Department of Physics Education, Universitas Sulawesi Barat, Indonesia[2]
Department of Physics Education, Universitas Negeri Yogyakarta, Indonesia[3]

purwokoharyadi.2021@student.uny.ac.id, purwokoharyadisantoso@unsulbar.ac.id

Abstract

Mechanical waves conceptual survey (MWCS) is a measurement tool established by the physics education research (PER) community to evaluate conceptual physics understanding of mechanical waves. A validation study is still needed to figure out the factor structure of MWCS using two data reduction techniques, namely exploratory factor analysis (EFA) and principal component analysis (PCA). The MWCS dataset in this paper was gathered from physics students ($n = 419$) from nineteen Ugandan secondary schools. The findings of this research suggested the single factor of the MWCS construct that has emerged from the dataset explored in this study. Several issues involved in the calculation of inter-item correlation within the dataset are suspected as the leading cause of the missing component solution or stable loading in the data. Moreover, there might be other issues that leave open space for future exploration. The findings reported in this paper could be the subject of further discussion in evaluating the validity of the MWCS as a research-based assessment (RBA) to measure students' conceptual understanding of wave mechanics within PER studies.

Keywords: principal component analysis, exploratory factor analysis, mechanical waves

Abstrak

Mechanical waves conceptual survey (MWCS) *adalah alat pengukuran dalam komunitas penelitian pendidikan fisika (PPF) untuk mengevaluasi pemahaman konsep fisika tentang gelombang mekanik. Studi validatsi masih perlu dilakukan dalam memahami struktur faktor MWCS melalui dua metode reduksi data yaitu* exploratory factor analysis (EFA) *dan* principal component analysis (PCA). *Dataset MWCS yang dianalisis dalam artikel ini dikumpulkan dari siswa fisika (n = 419) pada sembilan belas sekolah menengah di Uganda. Hasil studi menyarankan faktor tunggal dari konstruk MWCS teridentifikasi dari dataset yang dieksplorasi dalam penelitian ini. Beberapa masalah terlibat dalam korelasi antar butir diduga menjadi penyebab utama tidak ditemukannya solusi komponen atau loading yang stabil di dalam data. Selain itu, ada isu lain yang menyisakan ruang terbuka untuk eksplorasi lebih lanjut. Temuan yang dilaporkan dalam artikel ini masih bisa menjadi bahan diskusi lanjutan dalam mengevaluasi validitas dari MWCS sebagai sebuah instrumen penilaian berbasis penelitian (PBP) untuk mengukur pemahaman konsep siswa tentang gelombang mekanik dalam komunitas PPF.*

Kata kunci: principal component analysis, exploratory factor analysis*, gelombang mekanik*

Introduction

A mechanical wave is one of the physical topics discussed within the physics education curriculum from secondary school to college-level courses. A solid understanding of this topic is needed to study other physics concepts further. Therefore, the development of a conceptual inventory to measure students' understanding of this topic should be done either to measure the extent to which the wave concept can be grasped or to examine the effectiveness of the physics learning reforms.

Numerous works have studied how the mechanical wave has been understood (Bhathal et al., 2010; Caleon & Subramaniam, 2010; Eshach, 2014; Eshach & Schwartz, 2006; Hrepic et al., 2010; Kennedy & Laurillard, 2019; Kryjevskaia et al., 2011; Pejuan et al., 2012; Wittmann, 2002; Zeng et al., 2014). These previous studies contribute to identifying difficulties in students' understanding of mechanical waves. For instance, Wittmann (2010) mentions that students' approach to this topic could utilize object-oriented descriptions of wave pulses. Mechanical waves conceptual survey (MWCS) was first published by Tongchai et al. (2009). It is a conceptual survey on students' understanding of four constructs, namely propagation, superposition, reflection, and standing wave. To date, MWCS is still the most influential conceptual instrument on waves within the physics education research (PER) community compared to three other instruments reported by Barniol & Zavala (2016), Caleon & Subramaniam (2010), Hrepic et al. (2010), and Roedel et al. (1998).

Exploring the validity of a research-based assessment (RBA) should be administered routinely since the longer the PER community grows, the more studies should investigate the characteristics of their established RBA (Ding & Beichner, 2009). Formerly, the validity of the MWCS has been explored using classical item analysis, as reported by Tongchai et al. (2009). Further research with different analytical frameworks must have opportunities to conduct in evaluate the psychometric characteristics of MWCS. Evidence of validity should never cease to provide strong evidence that the MWCS is a valid instrument for a conceptual survey. It should be stated that MWCS can be employed as a sufficient RBA for the current development. However, a further study investigating the psychometric characteristics of MWCS should enable additional evidence of its current validity.

To enrich the value of MWCS, this paper is intended to explore the factor structure of MWCS through its pretest and post-test datasets using two data reduction techniques, namely principal component analysis (PCA) and exploratory factor analysis (EFA). Through PCA, we could learn how the structure of the analyzed dataset can form the factor structure that should belong to the MWCS measurement. Then, following the PCA and EFA, we will explain that the formed factors can classify the items in the set of factors. Nevertheless, in some literature, the definition and utilization of these techniques are often vague and interchangeable. There has been a long-lasting debate on the pros and cons between PCA and EFA, with some arguing in favor of the PCA technique (Steiger, 1979; Velicer & Jackson, 1990b, 1990a) and others in favor of the EFA technique (Bentler & Kano, 1990; Gorsuch, 2010; Mulaik, 1990). Distinctions between the two are primarily theoretical. As for research practices, the selection does not seem to affect empirical findings or substantive conclusions (Arrindell & van der Ende, 1985; Mulaik, 1990; Velicer et al., 1982). This study used both methods and discovered that these techniques are complementary rather than interchangeable. Insights reported in this paper can contribute two folds. First, it will expand the current evidence of the MWCS validity as a measurement tool disseminated by the PER community. Second, we see that the discussion about the difference between PCA and EFA still needs to be announced. Therefore, it can open broader insights to understand better these analytical frameworks used to analyze data in the validation studies.

Literature Review: Construct of the Mechanical Waves Conceptual Survey (MWCS)

The current version of the MWCS is designed based on the open-ended form developed by Wittmann (1998) through his dissertation project. From Wittmann (1998) point of view and other studies related to it, i.e., Wittmann et al. (1999), Wittmann (2002), and Wittmann et al. (2003), students' difficulties have been documented through several open-ended items. Nevertheless, multiple-choice items are inevitably the preferable format for administering the conceptual survey's large-scale setting. The MWCS is crafted as a multiple-choice test based on these previous investigations of students' difficulties with waves in four constructs. They are propagation, superposition, reflection, and standing waves. The original MWCS reported by Tongchai et al. (2009) discusses their test development phase and its standard psychometric analysis. MWCS was pilot tested on 632 Australian students (high school to second-year college students) and 270 Thai high school students.

Table 1 describes the original construct of the MWCS, initially proposed by Tongchai et al. (2009). The MWCS comprises 22 items, of which 17 items accommodate the conventional format of a multiple-choice test with different numbers of responses (five items with five responses, one item with three responses, five items with four responses, four items with six responses, and two items with eight responses). In addition to the multiple-choice (MC) format, MWCS also uses five two-tiered formats (TT), which assign students to perform their scientific reasoning within two stages of answer.

Table 1. Theoretical Description of MWCS's Construct.

Topic	Subtopic	No	Format	Note
Propagation	Sound variables	1	MC 4 options	
	Speed of sound waves	2	MC 4 options	
		3	MC 4 options	
	Speed of waves on strings	4	MC 6 options	
		5	MC 4 options	
	Displacement of medium in sound waves	6	MC 5 options	
		7	MC 8 options	Same option as in item 8
		8	MC 8 options	Same option as in item 8
Superposition	Construction	9	MC 6 options	
		10	MC 3 options	
	Destruction	11	MC 6 options	
		12	MC 4 options	
Reflection	Fixed end	13	MC 5 options	Same option as in item 14
		15	MC 5 options	
	Free end	14	MC 5 options	Same option as in item 13
		16	MC 5 options	
Sound waves	Transverse standing waves (strings)	17	TT 3 and 4 options	
		18	TT 3 and 4 options	
	Longitudinal standing waves (sound)	19	TT 3 and 4 options	
		20	MC 6 options	
		21	TT 3 and 5 options	
		22	TT 3 and 5 options	

In this paper, we decided to involve the original version of MWCS reported by Tongchai et al. (2009). Even Barniol & Zavala (2016) attempted the modified version of MWCS in 2016. We see the potential room for further exploring the factor structure of the original MWCS of Tongchai et al. (2009) that seems unknown from the provided literature within the community. However, we have to admit that this version can be further studied using another dataset from the modified version of MWCS.

To the best of our knowledge, Tongchai et al. (2009), Barniol & Zavala (2016), and most recently Kanyesigye et al. (2022) have just employed the classical analysis to conduct the validation studies of MWCS, including item difficulty, discrimination index, and internal reliability estimation. Instead of the classical theory's drawbacks in providing an understanding of the measurement quality, it is inevitably easier to conduct and interpret. To date, there is no doubt that most educational practices still consider it. Tongchai et al. (2009) reported that the original version of MWCS possessed a sufficient measure of item difficulty, discrimination index, and reliability estimation even though there is a little shift in several study contexts. The difficulty indices ranged between 0.2-0.8, as recommended by the conventional rule. Likewise, the discrimination indices are also evident to be acceptable to distinguish the higher and the lower students within a class. These are supplemented by the sufficient alpha Cronbach's reliability value estimated from 902 students as the sample.

One can argue that providing valid evidence of a measurement tool can be carried out with the perspective through which it has been administered. Comparing the evidence using other analytical frameworks should always be needed. Factor analysis is one of the psychometric analyses to examine how the constructs in Table 1 have been represented by the MWCS dataset (students) (Ding & Beichner, 2009; Hair et al., 2018). Through factor analysis, the factor structure of several RBAs within PER community has been understood to describe their validity of the conceptual physics measurement, i.e., *Force Concept Inventory* (FCI) (Eaton & Willoughby, 2018; Semak et al., 2017), *Force and Motion Conceptual Evaluation* (FMCE) (Wells et al., 2020), and *Colorado Learning Attitudes about Science Survey* (CLASS) (Kontro & Buschhüter, 2020). However, to the best of our inquiry, no evidence has been provided to determine the MWCS's construct validity. Using two reduction techniques reported in this paper, we can evaluate PCA and EFA findings to explore the item's ability to examine the structure of students' conceptual knowledge about mechanical waves.

Methods

The dataset analyzed in this paper was downloaded from Mendeley® cloud-based repository hosted as open data resources. Thus, everyone can access the data legally within the CC BY 4.0 license. The MWCS dataset was recently reported by Kanyesigye et al. (2022), a study conducted in southwest Uganda to evaluate problem-based learning in physics education. Data were gathered from February to April 2021 at the nineteen Ugandan secondary schools. *Mechanical Waves Conceptual Survey* (MWCS) (Tongchai et al., 2009), *Views about Science Survey* (VASS) (Halloun & Hestenes, 2002), and *Reformed Teaching Observation Protocol* (RTOP) (Sawada et al., 2002) were reported as a measure of the three variables probed in their study. However, for this paper, we determined to use the MWCS dataset among the available datasets reported by Kanyesigye et al. (2022). Interested readers are encouraged to involve those other datasets for attempting further studies.

Table 2. The Attribute of the MWCS Dataset Reported by Kanyesigye et al. (2022).

Students' attribute	Data type	Description
Class	Nominal	1 = experiment, 2 = control
Gender	Nominal	1 = female, 2 = male
Age	Ordinal	1 = < 16, 2 = 17, 3 = 18, 4 = 19, 5 = > 20
Major	Nominal	0 = none; 1 = Phys, Eco, Mat (PEM); 2 = Phys, Bio, Chem (PBC); 3 = Phys, Chem, Mat (PCM); 4 = Phys, Ent, Mat (PEM); 45 = Phys, Agri, Mat (PAM)
Status	Nominal	1 = female school, 2 = general
Owner	Nominal	1 = public, 2 = private

The dataset explored in this paper included students' scores in the MWCS pretest ($n = 239, \bar{x}_{\text{pre}} = 6.87, SD_{pre} = 2.38$) and the MWCS posttest ($n = 419, \bar{x}_{\text{post}} = 10.42, SD_{post} = 3.00$) separately. Table 2 demonstrates attributes, data types, and explanations of each attribute as a demographic aspect of the dataset that represents the context of the dataset. The different number of pretest and post-test data is driven by the different class assignments discussed in Kanyesigye et al. (2022).

The initial stage prior to the analytical process for researchers through secondary data sources is data preparation. The first thing to note was the case of missing data within Kanyesigye et al. (2022) dataset in several places. Fortunately, the missing value was less than the data size (< 5%). To address this potential issue, the imputation method with possible responses (Table 1) was chosen; hence the missing values were filled randomly (Kempf-Leonard, 2004). Subsequently, the student's score was coded based on the answer key provided in the dataset. Afterward, the dichotomous response was coded as one for correct and zero for the opposite. Then the results are filed on a separate sheet to simplify the working process. In addition, researchers also need to study the characteristics of the data to understand the context of the participants described in the dataset of Kanyesigye et al. (2022). The results are summarized in Table 3 about the demographic description of the MWCS pretest and post-test score regarding class, gender, age, major, school status, and school ownership.

Table 3. Demographic Description of MWCS Score on the Pretest and Post-Test Dataset.

Attribute	Group	MWCS pretest			MWCS posttest		
		n	$\bar{x}_{\text{pre}}$	SD_{pre}	n	$\bar{x}_{post}$	SD_{post}
Class	Experiment	132	6.44	2.52	231	11.00	3.13
	Control	102	7.39	2.09	188	9.72	2.67
Gender	Male	147	6.91	2.33	273	10.41	3.01
	Female	92	6.79	2.47	146	10.46	2.99
Age	< 16	4	6.75	1.5	6	10.17	2.93
	17	70	6.80	2.52	119	10.23	2.83
	18	114	6.64	2.35	182	10.88	3.11
	19	44	7.34	2.16	89	10.03	3.06
	> 20	7	8.29	2.81	23	9.34	2.35
Major	PEM	131	7.10	2.38	239	10.34	2.84
	PBC	22	7.36	1.79	38	9.86	2.29
	PCM	56	6.41	2.60	94	11.00	3.53
	PWM	14	5.85	2.18	19	10.79	3.85
	PAM	16	6.75	2.11	29	9.76	2.40
Status	Female	35	6.28	2.57	35	10.51	3.41
	General	204	6.70	2.34	384	10.42	2.97
Owner	Public	94	7.00	2.26	165	10.44	3.16
	Private	145	6.78	2.46	254	10.42	2.90

Exploring the factor structure of MWCS using principal component analysis (PCA) was carried out within R programming language employing several open-source packages such as "readxl" (Wickham et al., 2019), "psych" (Revelle et al., 2015), and "factoextra" (Kassambara et al., 2016). Naturally, these packages can be approached by every scholar in various fields, places, and any time. Packages "readxl" was used to import MWCS data filed as .xlsx format so that it can be translated within R data frame. Packages "psych" was a reliable library for psychometricians and can be easily identified by remembering the name "psych"ometry. These packages are the primary tool in PCA's work in this study. Then, "factoextra" was conducted in favor of the EFA technique supplemented with visualization plots so the reported results could be beautifully delivered.

After the running of PCA, EFA was then performed to examine the theoretical construct of MWCS (Table 1) in the same file of R scripts but was equipped with other packages complementing the PCA technique, i.e., "corpcor" (Schaefer et al., 2012), "ggplot2" (Wickham, 2017), "MASS" (Ripley et al., 2019), "MVN" (Korkmaz et al., 2014), and "psy" (Falissard, 2012). Some of these packages were supplementary because these packages are merely to provide summarized information or additional visualizations for a better readership of this paper.

Results and Discussion

Students' Response Patterns within the MWCS Pretest and Post-Test Dataset

The mean score of the MWCS pretest is 6.87 with a standard deviation of 2.38, and the MWCS post-test was 10.42 with a standard deviation of 3.00 of the 22 items examined in the MWCS. Remember that there are five "two-tier" questions (see Table 1). Students will be given a score if they can correctly answer the items and reasons. The normality distribution of pretest data can be evaluated from the skewness, less than two, and kurtosis, less than 7 (Kline, 2011). Based on this rule, it is found that both the MWCS pretest and post-test datasets are Gaussian distributed. Table 4 summarizes the students' responses on the MWCS pretest and post-test sessions, respectively.

Table 4. Students' Summary Response (as Proportion) of the MWCS Dataset. The Answer Keys are in Bold.

Topic	No	MWCS pretest (*n* = 239)							MWCS posttest (*n* = 419)						
		A	B	C	D	E	F	G	A	B	C	D	E	F	G
Propagation	1	.19	**.28**	.30	.23				.17	**.53**	.15	.14			
	2	**.15**	.30	.27	.28				**.49**	.15	.23	.13			
	3	.30	.24	**.23**	.23				.11	.14	**.58**	.16			
	4	.21	.19	.31	.26	.01	**.03**		.19	.09	.13	.10	.04	**.45**	
	5	**.19**	.27	.27	.27				**.46**	.25	.18	.11			
	6	.15	**.24**	.21	.27	.13			.07	**.51**	.17	.19	.06		
	7	.15	.26	**.12**	.22	.10	.07	.08	.08	.16	**.47**	.10	.06	.13	.01
	8	.14	.23	.21	.15	.14	**.06**	.07	.10	.08	.07	.17	.05	**.48**	.04
Superposition	9	.18	.31	.31	**.07**	.07	.05		.15	.16	.11	**.47**	.07	.04	
	10	**.15**	.57	.28					**.58**	.21	.21				
	11	.12	.25	**.20**	.29	.08	.06		.17	.08	**.47**	.17	.02	.09	
	12	**.15**	.38	.30	.17				**.56**	.16	.11	.17			
Reflection	13	.21	.21	.25	**.17**	.17			.13	.18	.11	**.47**	.10		
	15	.17	.15	.37	**.21**	.10			.19	.47	.15	**.10**	.10		
	14	.21	.10	.28	.23	**.18**			.10	.63	.12	.08	**.07**		
	16	.23	**.25**	.35	.11	.06			.50	**.12**	.15	.10	.13		
Sound waves	17	**.28**	.36	.36					**.22**	.57	.21				
	18	.28	**.34**	.37					.57	**.20**	.23				
	19	**.30**	.32	.38					**.20**	.50	.29				
	20	.19	**.35**	.19	.10	.11	.06		.18	**.17**	.14	.05	.05	.41	
	21	.28	.46	**.26**					.17	.32	**.51**				
	22	**.25**	.26	.49					**.25**	.57	.18				

In Table 4, we provide the proportion distribution of students' responses among the MWCS pretest and post-test datasets. Proportion is preferable since it would be more comparable to our discussion below. We treat both the pretest and post-test datasets as separate information to keep the nature of

different administrations. Conceptual changes can be discovered by the immediate shift of most selected responses between these two test sessions. For instance, most of the students (30%) answered "C" in the first MWCS pretest items. It is incorrect. Accordingly, most of them (53%) are successful in improving their conception of the correct answer ("B") in the post-test session. Nevertheless, eight MWCS items (14, 15, 16, 17, 18, 19, 20, and 22) were still discovered to be difficult to grasp by the students even though they attempted the post-test session.

Data Reduction of MWCS Dataset using PCA

Principal component analysis (PCA) parses the correlation matrix with the elements on the diagonal. The amount of variance is equal to the number of trace matrices or the number of observed variables contained in the analyzed data. PCA minimizes the sum of the distances perpendicular to the square of the component axis. The components in PCA analysis are uninterpretable. That is, there is no underlying construct. The principal component maintains the maximum amount of variance. The component score is a linear combination of the observed variables loaded by the eigenvectors.

Before PCA analysis can be performed through the MWCS pretest and post-test datasets, a statistical assumption must be examined, such as Bartlett's sphericity test. The null hypothesis declares that it is impossible to reduce the dimensions of the MWCS dataset. Table 5 reports the results of Bartlett's sphericity test. We see that the p-value is less than 0.05; thus, our null hypothesis is failed to reject. This value can articulate that PCA analysis could reduce the dimension of the MWCS dataset.

Table 5. Bartlett's Test of Sphericity.

MWCS pretest			MWCS pretest		
χ^2	p	df	χ^2	p	df
491.3805	0.000	351	149.905	0.000	351

Furthermore, the following fundamental thing that affects the overall data reduction results using PCA is the inter-items (variables) correlation matrix. Highly correlated items or variables are more likely to form components. Figure 1 shows that the MWCS pretest data cannot be categorized as highly correlated. In this picture, heatmap visualization with the circle representation is presented. We prefer to present using the circle visualization due to the more accessible interpretation. The "perfect" correlation is reported by the main diagonal of Figure 1. The way to read the heatmap visualization in Figure 1 is supported by the continuous spectrum of reference color put on the right side of the image. The color is spread as the spectrum of red and blue. The more red or blue, the correlation will be higher. However, if it is getting faded or even the same as the color of the paper on which this article is printed, then there is no significant correlation between the items. Figure 1 indicates that the MWCS dataset reported by Kanyesigye, et al. (2022) lacks correlation among the items. This may interfere with the overall results of the PCA, and even EFA discussed below. The visualization shown in Figure 1 is also found in the MWCS post-test data but with no substantive difference. Therefore, both pretest and post-test data have a sufficient inter-item correlation to form six factors, as in Table 1. We excluded the pretest correlation matrix to keep the length of this article still easy to read and worthy of publication because the results do not appear to be significantly different.
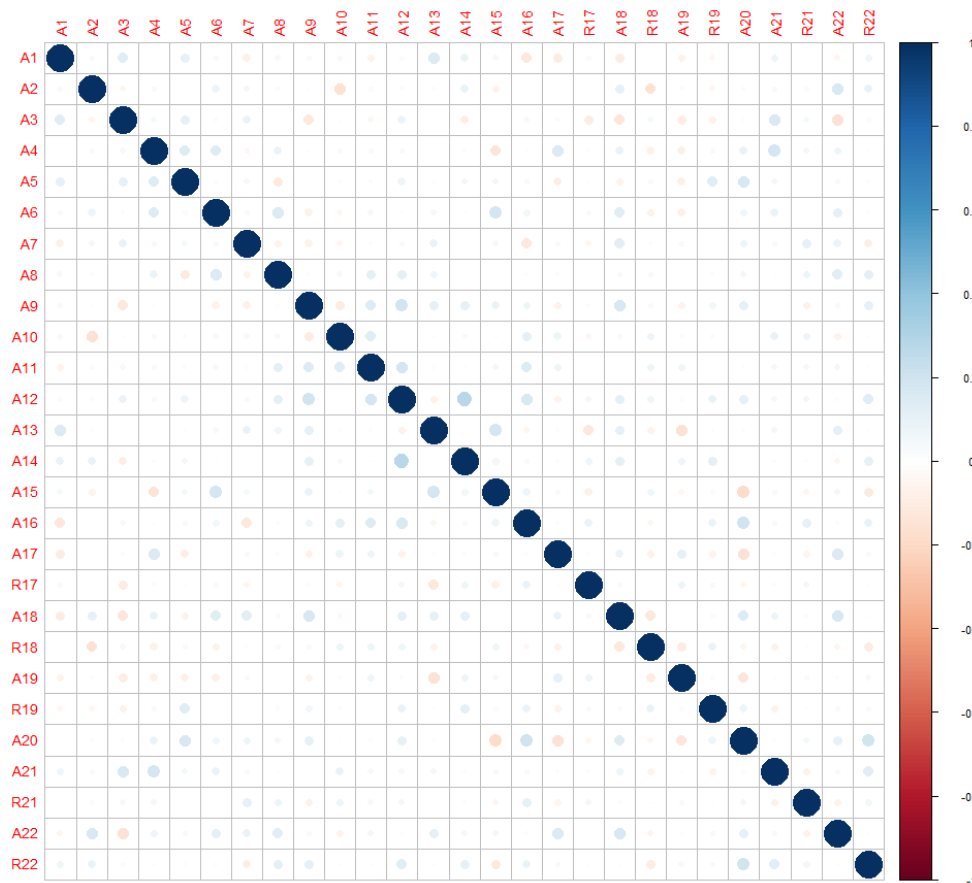
Figure 1. Heatmap Visualization of MWCS Pretest Correlation Matrix.

Table 6. Comparison of Eigen, Proportion of Variance, and Cumulative Proportion of PCA.

	MWCS pretest								
	prcomp			svd			princcomp		
	Eigen	Pr.V	Pr.C	Eigen	Pr.V	Pr.C	Eigen	Pr.V	Pr.C
PC1	1.40369	0.07298	0.07298	1.97033	0.07298	0.07298	1.89625	0.09925	0.09925
PC2	1.31113	0.06367	0.13664	1.71906	0.06367	0.13664	1.82882	0.09232	0.19157
PC3	1.27549	0.06025	0.19690	1.62687	0.06025	0.19690	1.65682	0.07577	0.26734
PC4	1.23590	0.05657	0.25347	1.52746	0.05657	0.25347	1.54872	0.06620	0.33354
PC5	1.21714	0.05487	0.30834	1.48142	0.05487	0.30834	1.46033	0.05886	0.39240
	MWCS post-test								
	prcomp			svd			princcomp		
	Eigen	Pr.V	Pr.C	Eigen	Pr.V	Pr.C	Eigen	Pr.V	Pr.C
PC1	1.888	0.132	0.132	3.5647	0.13203	0.13203	2.7299	0.1932	0.1932
PC2	1.2772	0.06042	0.19244	1.6312	0.06042	0.19244	1.8897	0.0926	0.2858
PC3	1.2592	0.05872	0.25116	1.5854	0.05872	0.25116	1.8056	0.0845	0.3703
PC4	1.2135	0.05454	0.30571	1.4726	0.05454	0.30571	1.5193	0.0598	0.4302
PC5	1.16947	0.05065	0.35636	1.3677	0.05065	0.35636	1.4876	0.0574	0.4875

PCA analysis then can be calculated in three ways, through the "prcomp" function, singular value decomposition or "svd", and "princcomp". The results through three different methods are reported in Table 6 and discover non-significant differences. In Table 6, five principal components (PCs) are only provided. We obtained 27 PCs within the MWCS dataset, which cannot be expected based on Table 1.

It can be underpinned by either "prcomp", "svd", or "princomp" reporting a small cumulative proportion (Pr.K) (less than fifty percent as recommended by the conventional rule). Due to the low value of Pr.K results, we cannot represent the students' understanding of mechanical waves due to a lack of acceptable measures. Rather, Alavi et al. (2020) argue that PCA has more emphasis on data reduction than interpretation. Therefore, the interpretation of these PCs should be noticed carefully, and we argue that the MWCS dataset tends to form a single-factor structure (see Figure 2 dan 3 below). Unsurprisingly, this can indicate that the MWCS dataset is less representative of validating theoretical construct as described in Table 1.
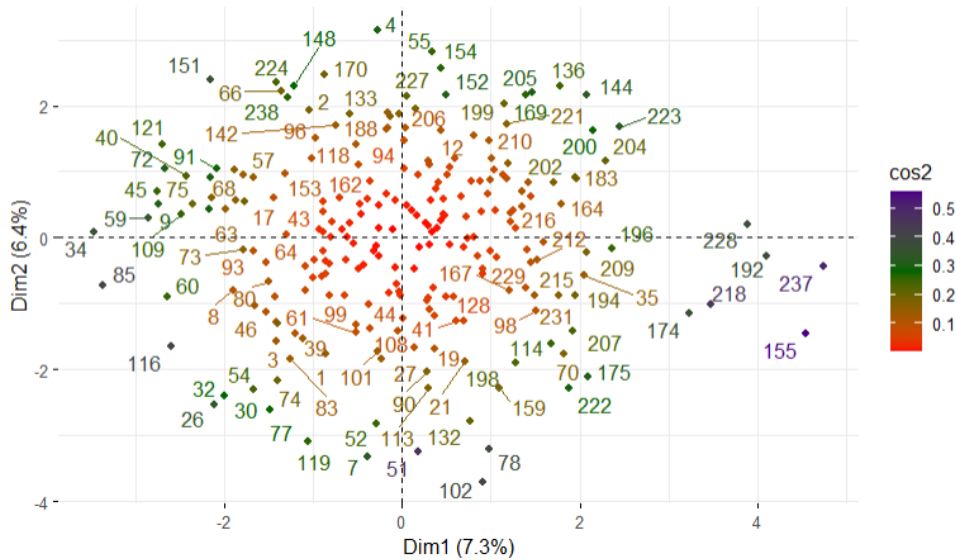
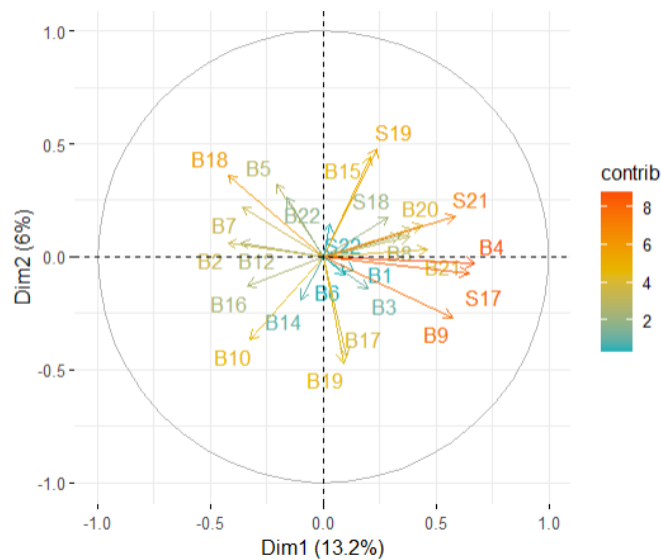**Figure 2**. MWCS Pretest Individual (Respondent) Curves (Individual Group)

Figure 3. PCA Visualization in the Perspective of the MWCS Post-Test Variables.

Some visualizations presented by the "factoextra" packages are demonstrated in Figure 2 (by individuals/ students) and Figure 3 (by variables/ items). We can elaborate that those results tend to be clustered as one group. They make the distinctions among the individuals/ students troublesome to identify. In this figure, the expected grouped representation should be clustered within four distinct

theoretical regions, as in Table 1. Instead, we can suspect this immediate single-factor pattern can be driven by the former lack of an inter-item correlation matrix (Figure 1).

Furthermore, we can group by variables (items) in Figure 3. Naturally, this must be what EFA does in the next part of our analysis. Accordingly, the single dimension is also demonstrated within the aggregation of variables. This figure essentially makes clustering emerge in the structure of MWCS items. The variance explained by the emerged dimension is neither expected from the first nor second dimensions (less than fifty percent as recommended by the conventional rule). Hence, interpreting these dimensions to examine the theoretical construct of Tongchai et al. (2009) can be problematic to be informed for wider contexts.

Figure 4 describes a scree plot describing the number of PCA emerging components. This figure suggests that there is merely one factor represented by the MWCS dataset based on the "elbow" method. After the first component Figure 4 makes a substantive "elbow" in the second component, and it leads to the conclusion that the MWCS dataset reported by Kanyesigye et al. (2022) can be problematic in providing empirical validity to the theoretical foundations (Table 1). Even though we can confirm that the percentage obtained in the y-axes is less than fifty percent as the conventional rule recommends. Therefore, our PCA result can not represent the MWCS dataset sufficiently.
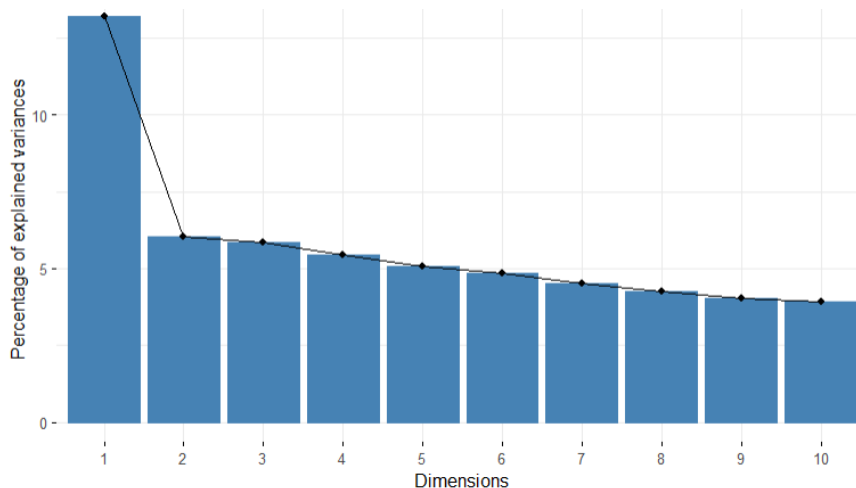

Figure 4. Scree Plot of Variance Explained by MWCS Post-Test Components

Exploratory Factor Analysis (EFA) and MWCS Dataset

Factor analysis approaches data reduction from a fundamentally different perspective from the PCA. Factor analysis is a measurement model of latent variables. The latent construct is unable to be directly measured by a single variable, i.e., students' understanding of mechanical waves in our case. Latent variables are measured through causal relationships within the MWCS items. In this section, we will report the EFA results of the MWCS dataset using two extraction methods: principal component analysis (PCA) and principal axis factoring (PAF). The same "PCA" terminology can be misleading with the previous analysis as if they share standard features. They must be distinct theoretically. Two different rotation methods of EFA are also examined, both "varimax" from the orthogonal matrix and "oblimin" from the diagonal matrix.

Several steps must be examined prior to the factor analysis. Some have been reported previously at the PCA stage. Hence they will not be discussed here again. The different statistical tests of EFA are the need for a measure of sampling adequacy (MSA) test through the Kaiser-Meyer-Olkin (KMO) test and the size of the anti-image correlation matrix. Both the MWCS pretest and post-test satisfy the KMO cutoff

value higher than 0.5, and the diagonal value of the anti-image correlation matrix is more significant than 0.5. Therefore, we can confidently conclude that the MWCS dataset can be reduced as its factor structure with EFA.

To extract the constructs explained by the MWCS dataset, EFA also uses PCA as its extraction method. However, this extraction is different from what has been executed formerly. PCA assumes no outliers in the data, while EFA assumes a multivariate normal distribution in the maximum likelihood extraction method. In contrast to PCA, EFA analysis decomposes the "adjusted" correlation matrix. The diagonals of the correlation matrix have been adjusted for unique factors. The amount of variance explained is the same as the number of trace matrices or commonalities. Factors represent the expected variance in the dataset. Squared multiple correlations (SMC) were defined to estimate commonality on the diagonal. The observed variable is a linear combination of unique and fundamental factors.

EFA is generally purposed to investigate the construct's empirical structure and estimate the instrument's reliability. EFA is recommended for researchers who do not have a hypothesis about the instrument's structure (Kline, 2011). Several extraction methods are used to determine the number of factors that will be formed. They are principal component analysis (PCA), principal axis factoring (PAF), maximum likelihood (canonical factoring), alpha factoring, and image factoring. In this paper, authors decide to perform two extraction methods most often chosen in several previous studies (Eaton & Willoughby, 2018; Scott et al., 2012; Semak et al., 2017). This method tries to determine the smallest number of factors that can represent the variability of the original variable, which is related to the factors (this contrasts with the PCA we did earlier, which looked for factors that reflect the variability of the variables). Both PCA and EFA should have relevant results if the variables are highly correlated and/or the number of original variables is very large.

Table 7. Comparison of the EFA Results using Two Different Extraction Methods (PCA and PAF).

	MWCS pretest							
	PCA (mean complex = 2.2)				**PAF factoring (mean complex = 2.1)**			
	PC1	PC2	PC3	PC4	PA1	PA2	PA3	PA4
SS loading	1.97	1.72	1.63	1.53	1.16	0.88	0.78	0.72
Pr. Var	0.07	0.06	0.06	0.06	0.04	0.03	0.03	0.03
Cum. V	0.07	0.14	0.20	0.25	0.04	0.08	0.10	0.13
Pr. Exp	0.29	0.25	0.24	0.22	0.33	0.25	0.22	0.20
Cum. P	0.29	0.54	0.78	1.00	0.33	0.58	0.80	1.00
	MWCS post-test							
	PCA (mean complex = 2.2)				**PAF factoring (mean complex 2.1)**			
	PC1	PC2	PC3	PC4	PA1	PA2	PA3	PA4
SS loading	3.56	1.63	1.59	1.47	2.86	0.87	0.78	0.69
Pr. Var	0.13	0.06	0.06	0.05	0.11	0.03	0.03	0.03
Cum. V	0.13	0.19	0.25	0.31	0.11	0.14	0.17	0.19
Pr. Exp	0.43	0.20	0.19	0.18	0.55	0.17	0.15	0.13
Cum. P	0.43	0.63	0.82	1.00	0.55	0.72	0.87	1.00

Four principal components (PCs) are reported in Table 7 as a selection of the number of factors extracted from the EFA technique. This number of PCs was determined based on the theoretical construct proposed by Tongchai, et al. (2009) in Table 1. In Table 6 above, five PCs are reported. Contrarily, they are theoretically distinct. In Table 7, the number of factors extracted by the EFA method should be determined prior to the calculation. Our determination was made based on the theoretical construct in Table 1 though we must admit that EFA underlies the assumption of exploratory than confirmatory. Then, we can evaluate the results by considering the Cum. V (an abbreviation of

cumulative variance) is discussed in Table 6 above. We can infer from Table 7 that four components cannot explain the variance of the MWCS dataset (less than 50% as recommended by the literature). In summary, our findings suggest that the MWCS dataset possesses a single factor and fails to form four constructs as expected theoretically.

Table 7 also reveals the results of two different extraction methods, PCA and PAF. We use the terminology of "PC" for PCA results and "PA" (an abbreviation of the principal axis) for PAF extraction method. Contrasting both discover common features in some places, and we can discover no significant difference in the mean item complexity. The "SS loadings" row is the sum of squared loadings. This is sometimes used to determine the value of a particular factor. We say a factor is worth keeping if the SS loading is greater than 1. In our example, PCA results possess values greater than 1 and just one "PA" which satisfies the cutoff value. The "Pr. Var" row reflects the proportion of variance explained by a particular factor. We can discover a small variance explained by our four PCs and PAs.

Table 8. Comparison of Two Different EFA Rotation Methods (Varimax and Oblimin).

| | MWCS pretest | | | | | | | |
| | Varimax (mean complex = 1.8) | | | | Oblimin (mean complex = 1.9) | | | |
	PC1	PC2	PC3	PC4	PC1	PC2	PC3	PC4
SS loading	0.99	0.96	0.81	0.79	1.02	0.98	0.78	0.76
Pr. Var	0.04	0.04	0.03	0.03	0.04	0.04	0.03	0.03
Cum. V	0.04	0.07	0.10	0.13	0.04	0.07	0.10	0.13
Pr. Exp	0.28	0.27	0.23	0.22	0.29	0.28	0.22	0.21
Cum. P	0.28	0.55	0.78	1.00	0.29	0.56	0.79	1.00
	MWCS post-test							
	Varimax (mean complex = 1.8)				Oblimin (mean complex = 2)			
	PC1	PC2	PC3	PC4	PC1	PC2	PC3	PC4
SS loading	2.22	1.20	0.94	0.83	2.14	1.29	1.01	0.76
Pr. V	0.08	0.04	0.03	0.03	0.08	0.05	0.04	0.03
V. K	0.08	0.13	0.16	0.19	0.08	0.13	0.16	0.19
Pr. D	0.43	0.23	0.18	0.16	0.41	0.25	0.19	0.15
Pr. K	0.43	0.66	0.84	1.00	0.41	0.66	0.85	1.00

In Table 8, the varimax and oblimin extraction method find the same mean item complexity though little difference occurs at the oblimin of the MWCS post-test. It seems that oblimin method obtains lower non-significant results than the varimax in all aspects (SS loading, Pr. Var, Cumulative Variance, Pr. Exp, and Cumulative Proportion). These findings align with the recommendation that both EFA and CFA frameworks could erroneously be used to analyze the same data and even yield similar results (Yanai & Ichikawa, 2006).

Internal Reliability of MWCS Dataset (*Cronbach's α*)

This study calculated Cronbach's alpha measure to estimate the reliability of the MWCS instrument. Neither pretest nor post-test MWCS can be concluded as adequate reliability ($\alpha_{pre} = 0.352, \alpha_{post} = 0.284$). This value is unsatisfactory and can be a measure that the factor structure solution generated in this article is unclear to be generalized in the broader context. Internal consistency, such as Cronbach's alpha coefficient, highly depends on the total item correlation. The inter-item correlation also indirectly affects the total item correlation (Figure 2). It can be seen in Figure 2 that the inter-item correlation is problematic in the Kanyesigye et al. (2022). In addition, the opportunity for discrepancies or student misconceptions in the sample is entirely possible.

Exploring MWCS Construct using PCA and EFA Dataset Reported by Kanyesigye et al. (2022)

This study aims to extract the components or factors in the MWCS dataset to understand the MWCS construct proposed by Tongchai et al. (2009) using the dataset reported by Kanyesigye et al. (2022) empirically. Before we discuss it further, we can start this discussion from the descriptive distribution of the MWCS pretest and post-test responses in Table 4 above. The numbers in bold are the answer keys that most respondents should answer. There is a conceptual shift in the respondent group from the pretest to the post-test. This can be explained rationally because the context of the dataset is a problem-based learning experiment, but this effect is outside the purpose of this study. Starting with the first item, the respondents were evenly distributed throughout the answers (A-D) at the pretest and tended to choose option B at the post-test. This result implies that item 1 means working well. However, some items fail to be good items.

One can check the items where the participants are not dominated by the correct answers in the fifteenth item. At the post-test, students are congregated on option B (incorrect answer). Logical reasons can be conveyed within two folds of possibilities. First, this item failed to be understood by the students; hence those with high abilities even took the wrong answer. Failure to understand the items can be driven by grammatical errors or the confusing effect of distractors. In addition, this finding can be related to the students' misconceptions at the class level, and physics teachers must evaluate their instruction. Misconceptions usually occur in minority groups in the classroom. Misconceptions in large groups indicate incorrect content delivery or even item construction errors.

PCA results should be able to describe the number of components that can be reduced from the MWCS dataset. The results of the variance components that can be explained from the dataset have been summarized in Table 6 with three applied methods. Some literature suggests that the total variance explained by all components should be between 70-80 percent of the item variance. However, there is another argument that social studies can use the cutoff of 50-60 percent as acceptable results for inferring the final represented components. However, if we examine the results in column Pr. K (cumulative proportion) found that the result is inadequate. Extracted components based on the ideas proposed by Tongchai et al. (2009) are expected as four topics in the MWCS theoretically (propagation, superposition, reflection, and sound waves) (Table 2). The MWCS dataset provided by Kanyesigye et al. (2022) already fulfilled several analytical assumptions, such as data normality, Barttlet sphericity test, and KMO sample adequacy test. However, as previously explained, our dataset is somewhat problematic in the basic aspect of factor analysis, namely the inter-item correlation matrix (Figure 2). The correlation problem is also related to the reliability value not suggested by the statistical literature and requires an alpha value greater than 0.7. These issues allow other researchers to examine these data using more advanced analysis, such as item response theory (Hansen & Stewart, 2021; Smith et al., 2020; Stewart et al., 2018).

Several extractions and rotation methods of EFA have been approached in Table 8. Although showing somewhat different results from previous PCA (not extraction method within EFA), the PCA extraction method still shows a cumulative proportion that is less than the minimum suggested by the literature and is not expected from the findings of Tongchai et al. (2009). We can declare that PCA and EFA found similar results to conclude the factor structure of the MWCS dataset. Nevertheless, EFA cannot replace the merit of PCA as a data reduction technique. Both have different theoretical frameworks, and they are complementary.

We discover that the theoretical construct of the MWCS cannot be empirically satisfied based on Kanyesigye, et al. (2022) dataset. Several issues can drive this finding. One fundamental factor is the inadequate correlated items within the MWCS pretest and post-test dataset. Furthermore, one can argue that respondents need to be increased for cases with low correlation(Schreiber, 2021). This case seems to contradict the KMO sample adequacy test, which has claimed that the data used in this study is

significantly sufficient (Table 5). However, the next question is whether larger data points with the same instrument guarantee MWCS reliability. This question can probably be answered intuitively. According to Schreiber (2021), before deciding to collect more data, we should return to the nature of the measurement itself. Valid data support valid measurements. Valid data will be gathered from the good items. The items mentioned above are still problematic in MWCS, which several aspects can influence. As explained earlier, some items do not work correctly. Further research opportunities should be planned by diagnosing them through more advanced analysis.

Conclusion

Investigating the psychometric characteristics of research-based assessment (RBA) is an ongoing process of validity studies for better measurement results. In this paper, MWCS, as one of the conceptual inventories within the PER community, has been examined for its construct validity using two frameworks, PCA and EFA. Nevertheless, we discover no substantial evidence to empirically conclude the established theoretical factor structure of MWCS via these frameworks. We suspect that the most fundamental factor directly affecting our PCA and EFA findings is the inadequate inter-item correlation matrix that emerged from the employed dataset. The low value can contribute to the unstable solution for components or factors extracted by the PCA and EFA.

We highlight that PCA and EFA are distinct, and these techniques cannot be interchangeable but complementary. We suggest that the misinterpretation in favor of these distinct analytical frameworks should be better consulted theoretically. On the other hand, findings from our studies can be valuable insight for extending the current milestone to provide evidence of the MWCS validity. There is room for further study to see whether the research context influences the factor structure measured by the MWCS. The constructs extracted from this study still need to be examined further. Future researchers are encouraged to review the construction of MWCS items and evaluate the format that must be administered. Several recommendations for better conducting and interpreting the PCA and EFA results are provided in our paper.

References

Alavi, M., Visentin, D. C., Thapa, D. K., Hunt, G. E., Watson, R., & Cleary, M. (2020). Exploratory factor analysis and principal component analysis in clinical studies: Which one should you use? In *Journal of Advanced Nursing*. https://doi.org/10.1111/jan.14377

Arrindell, W. A., & van der Ende, J. (1985). An empirical test of the utility of the observations-to-variables ratio in factor and components analysis. *Applied Psychological Measurement*, *9*(2). https://doi.org/10.1177/014662168500900205

Barniol, P., & Zavala, G. (2016). Mechanical waves conceptual survey: Its modification and conversion to a standard multiple-choice test. *Physical Review Physics Education Research*, *12*(1). https://doi.org/10.1103/PhysRevPhysEducRes.12.010107

Bentler, P. M., & Kano, Y. (1990). On the equivalence of factors and components. *Multivariate Behavioral Research*, *25*(1). https://doi.org/10.1207/s15327906mbr2501_8

Bhathal, R., Sharma, M. D., & Mendez, A. (2010). Educational analysis of a first-year engineering physics experiment on standing waves: Based on the ACELL approach. *European Journal of Physics*, *31*(1). https://doi.org/10.1088/0143-0807/31/1/003

Caleon, I., & Subramaniam, R. (2010). Development and application of a three-tier diagnostic test to assess secondary students' understanding of waves. *International Journal of Science Education*, *32*(7). https://doi.org/10.1080/09500690902890130

Ding, L., & Beichner, R. (2009). Approaches to data analysis of multiple-choice questions. *Physical Review Special Topics - Physics Education Research*, *5*(2). https://doi.org/10.1103/PhysRevSTPER.5.020103

Eaton, P., & Willoughby, S. D. (2018). Confirmatory factor analysis applied to the force concept inventory. *Physical Review Physics Education Research*, *14*(1). https://doi.org/10.1103/PhysRevPhysEducRes.14.010124

Eshach, H. (2014). Development of a student-centered instrument to assess middle school students conceptual understanding of sound. *Physical Review Special Topics - Physics Education Research*, *10*(1). https://doi.org/10.1103/PhysRevSTPER.10.010102

Eshach, H., & Schwartz, J. L. (2006). Sound stuff? Naïve materialism in middle-school students' conceptions of sound. *International Journal of Science Education*, *28*(7). https://doi.org/10.1080/09500690500277938

Falissard, B. (2012). Package 'psy'. Various procedures used in psychometry version. *R Package Version 1.1, 20*.

Gorsuch, R. L. (2010). Common factor analysis versus component analysis: some well and little known facts. *Http://Dx.Doi.Org/10.1207/S15327906mbr2501_3*, *25*(1), 33–39. https://doi.org/10.1207/S15327906MBR2501_3

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2018). *Multivariate data analysis*. Pearson. https://doi.org/10.1016/j.ijpharm.2011.02.019

Halloun, I., & Hestenes, D. (2002). *Interpreting VASS dimensions and profiles for physics students*.

Hansen, J., & Stewart, J. (2021). Multidimensional item response theory and the brief electricity and magnetism assessment. *Physical Review Physics Education Research*, *17*(2), 020139. https://doi.org/10.1103/PHYSREVPHYSEDUCRES.17.020139/FIGURES/3/MEDIUM

Hrepic, Z., Zollman, D. A., & Rebello, N. S. (2010). Identifying students' mental models of sound propagation: The role of conceptual blending in understanding conceptual change. *Physical Review Special Topics - Physics Education Research*, *6*(2). https://doi.org/10.1103/PhysRevSTPER.6.020114

Kanyesigye, S. T., Uwamahoro, J., & Kemeza, I. (2022). Data collected to measure the impact of problem-based learning and document physics classroom practices among Ugandan secondary schools. *Data in Brief*, *44*, 108534. https://doi.org/10.1016/J.DIB.2022.108534

Kassambara, A., of, F. M.-E. and visualize the results, & 2017, undefined. (2016). Package "factoextra." *Cran.Microsoft.Com*. https://cran.microsoft.com/snapshot/2016-11-30/web/packages/factoextra/factoextra.pdf

Kempf-Leonard, K. (2004). Encyclopedia of social measurement. In *Encyclopedia of Social Measurement*. https://doi.org/10.5860/choice.42-5629

Kennedy, E., & Laurillard, D. (2019). The potential of MOOCs for large-scale teacher professional development in contexts of mass displacement. *London Review of Education*, *17*(2). https://doi.org/10.18546/LRE.17.2.04

Kline, R. B. (2011). *Principles and practice of structural equation modeling, fourth edition - Rex B. Kline - Google Books*. 9.

Kontro, I., & Buschhüter, D. (2020). Validity of Colorado learning attitudes about science survey for a high-achieving, Finnish population. *Physical Review Physics Education Research*, *16*(2). https://doi.org/10.1103/PhysRevPhysEducRes.16.020104

Korkmaz, S., Goksuluk, D., & Zararsiz, G. (2014). MVN: An R package for assessing multivariate normality. *R Journal*, *6*(2). https://doi.org/10.32614/rj-2014-031

Kryjevskaia, M., Stetzer, M. R., & Heron, P. R. L. (2011). Student understanding of wave behavior at a boundary: The limiting case of reflection at fixed and free ends. *American Journal of Physics*, *79*(5). https://doi.org/10.1119/1.3560430

Mulaik, S. A. (1990). Blurring the distinctions between component analysis and common factor analysis. *Multivariate Behavioral Research*, *25*(1). https://doi.org/10.1207/s15327906mbr2501_6

Pejuan, A., Bohigas, X., Jaén, X., & Periago, C. (2012). Misconceptions about sound among engineering students. *Journal of Science Education and Technology*, *21*(6). https://doi.org/10.1007/s10956-011-9356-6

Revelle, W., network, M. R.-T. comprehensive R. archive, & 2015, undefined. (2015). Package "psych." *Cran.Rstudio.Org*. https://cran.rstudio.org/web/packages/psych/psych.pdf

Ripley, B., Venables, B., Bates, D. M., Hornik, K., Gebhardt, A., & Firth, D. (2019). Package 'MASS' (Version 7.3-51.4). *Cran-R Project*.

Roedel, R. J., El-Ghazaly, S., Rhoads, T. R., & El-Sharawy, E. (1998). Wave concepts inventory - an assessment tool for courses in electromagnetic engineering. *Proceedings - Frontiers in Education Conference*, *2*. https://doi.org/10.1109/FIE.1998.738761

Sawada, D., Piburn, M. D., Judson, E., Turley, J., Falconer, K., Benford, R., & Bloom, I. (2002). Measuring reform practices in science and mathematics classrooms: the reformed teaching observation protocol. *School Science and Mathematics*, *102*(6), 245–253. https://doi.org/10.1111/J.1949-8594.2002.TB17883.X

Schaefer, A. J., Opgen-rhein, R., Zuber, V., Pedro, A., Silva, D., & Strimmer, K. (2012). Package ' corpcor .' *Graphical Models*.

Schreiber, J. B. (2021). Issues and recommendations for exploratory factor analysis and principal component analysis. *Research in Social and Administrative Pharmacy*, *17*(5), 1004–1011. https://doi.org/10.1016/J.SAPHARM.2020.07.027

Scott, T. F., Schumayer, D., & Gray, A. R. (2012). Exploratory factor analysis of a force concept inventory data set. *Physical Review Special Topics - Physics Education Research*, *8*(2). https://doi.org/10.1103/PhysRevSTPER.8.020105

Semak, M. R., Dietz, R. D., Pearson, R. H., & Willis, C. W. (2017). Examining evolving performance on the Force Concept Inventory using factor analysis. *Physical Review Physics Education Research*, *13*(1). https://doi.org/10.1103/PhysRevPhysEducRes.13.010103

Smith, T. I., Louis, K. J., Ricci, B. J., & Bendjilali, N. (2020). Quantitatively ranking incorrect responses to multiple-choice questions using item response theory. *Physical Review Physics Education Research*, *16*(1). https://doi.org/10.1103/PhysRevPhysEducRes.16.010107

Steiger, J. H. (1979). Factor indeterminacy in the 1930's and the 1970's some interesting parallels. *Psychometrika*, *44*(2). https://doi.org/10.1007/BF02293967

Stewart, J., Zabriskie, C., Devore, S., & Stewart, G. (2018). Multidimensional item response theory and the Force Concept Inventory. *Physical Review Physics Education Research*, *14*(1). https://doi.org/10.1103/PhysRevPhysEducRes.14.010137

Tongchai, A., Sharma, M. D., Johnston, I. D., Arayathanitkul, K., & Soankwan, C. (2009). Developing, evaluating and demonstrating the use of a conceptual survey in mechanical waves. *International Journal of Science Education*, *31*(18). https://doi.org/10.1080/09500690802389605

Velicer, W. F., & Jackson, D. N. (1990a). Component analysis versus common factor analysis: some further observations. *Multivariate Behavioral Research*, *25*(1). https://doi.org/10.1207/s15327906mbr2501_12

Velicer, W. F., & Jackson, D. N. (1990b). Component analysis versus common factor analysis: some issues in selecting an appropriate procedure. *Multivariate Behavioral Research*, *25*(1). https://doi.org/10.1207/s15327906mbr2501_1

Velicer, W. F., Peacock, A. C., & Jackson, D. N. (1982). A comparison of component and factor patterns: A monte carlo approach. *Multivariate Behavioral Research*, *17*(3). https://doi.org/10.1207/s15327906mbr1703_5

Wells, J., Henderson, R., Traxler, A., Miller, P., & Stewart, J. (2020). Exploring the structure of misconceptions in the force and motion conceptual evaluation with modified module analysis. *Physical Review Physics Education Research*, *16*(1). https://doi.org/10.1103/PHYSREVPHYSEDUCRES.16.010121

Wickham, H. (2017). ggplot2 - Elegant graphics for data analysis (2nd Edition). In *Journal of Statistical Software* (Vol. 77, Issue Book Review 2).

Wickham, H., Bryan, J., Kalicinski, M., & Valery, K. (2019). *Package "readxl."* https://cran.microsoft.com/snapshot/2019-03-09/web/packages/readxl/readxl.pdf

Wittmann, M. C. (1998). Making sense of how students come to an understanding of physics: An example from mechanical waves. *Physics*.

Wittmann, M. C. (2002). The object coordination class applied to wave pulses: analyzing student reasoning in wave physics. *Http://Dx.Doi.Org/10.1080/09500690110066944*, *24*(1), 97–118. https://doi.org/10.1080/09500690110066944

Wittmann, M. C., Steinberg, R. N., & Redish, E. F. (1999). Making sense of how students make sense of mechanical waves. *The Physics Teacher*, *37*(1). https://doi.org/10.1119/1.880142

Wittmann, M. C., Steinberg, R. N., & Redish, E. F. (2003). Understanding and affecting student reasoning about sound waves. *International Journal of Science Education*, *25*(8). https://doi.org/10.1080/09500690305024

Yanai, H., & Ichikawa, M. (2006). Factor analysis. *Handbook of Statistics*, *26*, 257–296. https://doi.org/10.1016/S0169-7161(06)26009-7

Zeng, L., Smith, C., Poelzer, G. H., Rodriguez, J., Corpuz, E., & Yanev, G. (2014). Illustrations and supporting texts for sound standing waves of air columns in pipes in introductory physics textbooks. *Physical Review Special Topics - Physics Education Research*, *10*(2). https://doi.org/10.1103/PhysRevSTPER.10.020110