

Cogent Psychology

Country United Kingdom -  [SJR Ranking of United Kingdom](#)

Subject Area and Category [Psychology](#)
[Psychology \(miscellaneous\)](#)

Publisher [Cogent OA](#)

Publication type Journals

ISSN 23311908

Coverage 2014-2020

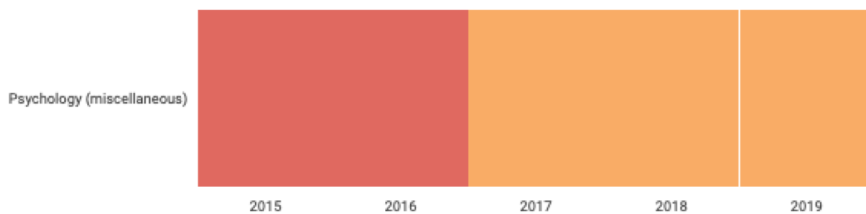
11

H Index

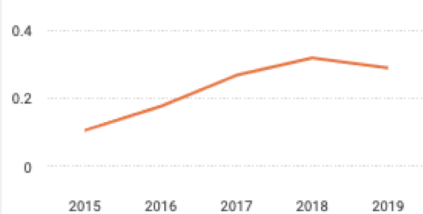
Scope One of the largest multidisciplinary open access journals serving the psychology community, Cogent Psychology provides a home for scientifically sound peer-reviewed research. Part of Taylor & Francis / Routledge, the journal provides authors with fast peer review and publication and, through open access publishing, endeavours to help authors share their knowledge with the world. Cogent Psychology particularly encourages interdisciplinary studies and also accepts replication studies and negative results. Cogent Psychology covers a broad range of topics and welcomes submissions in all areas of psychology, ranging from social psychology to neuroscience, and everything in between. Led by Editor-in-Chief Professor Peter Walla of Webster Private University, Austria, and supported by an expert editorial team from institutions across the globe, Cogent Psychology provides our authors with comprehensive and quality peer review. Rather than accepting manuscripts based on their level of importance or impact, editors assess manuscripts objectively, accepting valid, scientific research with sound rigorous methodology. Article-level metrics let the research speak for itself.

-  [Homepage](#)
- [How to publish in this journal](#)
- [Contact](#)
-  [Join the conversation about this journal](#)

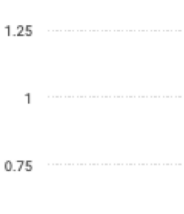
Quartiles



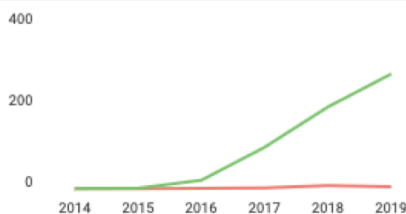
SJR



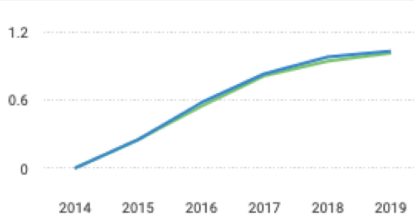
Citations per document



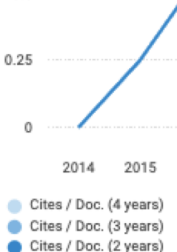
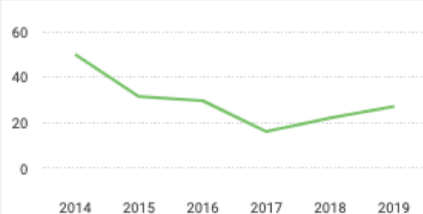
Total Cites



External Cites per Doc



% International Collaboration







Journal

Cogent Psychology >

[Submit an article](#)

[Journal homepage](#)

[New content alerts](#)

[RSS](#)

[Citation search](#)

[Current issue](#) [Browse list of issues](#)

About this journal

[Aims and scope](#)

[Instructions for authors](#)

[Journal information](#)

[Editorial board](#)

[Advertising information](#)



[Editorial policies](#)

Editorial board

Editor-in-Chief:

Peter Walla (Newcastle University, Australia)

Senior Editors:

Stefan Elmer (University of Zurich, Switzerland)

Editors:

Jubin Abutalebi (Vita-Salute San Raffaele University, Italy)

Arjen Alink (University Medical Center Hamburg-Eppendorf, Germany)

Asghar Arfaie, MD (Tabriz University of Medical Sciences, Iran)

Michela Balsamo (University of Chieti-Pescara, Italy)

Jens Binder (Nottingham Trent University, UK)

Kulbir Birak (University Campus Suffolk, UK)

Mathilde Bonnefond (Radboud University Nijmegen, Netherlands)

Delane Botelho (Fundacao Getulio Vargas, Brazil)

Josef Brockmann (Free Praxis, Germany)

David R. Brodbeck (Algoma University, Canada)

[Submit an article](#)

[New content alerts](#)

[RSS](#)

[Citation search](#)

[Current issue](#) [Browse list of issues](#) [Explore](#)

Rizwana Roomaney & Ashley Lourens | Carmen Rodriguez-Blazquez (Reviewing editor)

Article: 1823608

Published online: 29 Sep 2020

[Abstract](#) | [Full Text](#) | [References](#) | [PDF \(362 KB\)](#)

0
CrossRef citations

0
Altmetric

APPLIED PSYCHOLOGY | REVIEW ARTICLE

[Is long-term ABA therapy abusive: A response to Sandoval-Norton and Shkedy](#) >

Kathryn A. Gorycki, Paula R. Ruppel & Thomas Zane | Carryl P. Navalta (Reviewing editor)

Article: 1823615

Published online: 23 Sep 2020

[Abstract](#) | [Full Text](#) | [References](#) | [PDF \(376 KB\)](#)

391
Views

0
CrossRef citations

0
Altmetric

APPLIED PSYCHOLOGY | RESEARCH ARTICLE

[Psychometric assessment of an instrument evaluating the effects of affective variables on students' WTC in face-to-face and digital environment](#) >

Herri Mulyono, Regitha Saskia, Vista Sulaim Arrummaiza & Gunawan Suryoputro | Lucia Monacis (Reviewing editor)

Article: 1823617

Published online: 23 Sep 2020

[Abstract](#) | [Full Text](#) | [References](#) | [PDF \(447 KB\)](#)

146
Views

0
CrossRef citations

0
Altmetric

[Job seeker attraction to organizational justice mediated by organizational reputation](#) >

Hasnun Anip Bustaman, Mohammad Nazri Mohd nor, Azni Zarina Taha & Maheran Zakaria | Giulia Prete (Reviewing editor)

Article: 1816255

Published online: 10 Sep 2020

[Abstract](#) | [Full Text](#) | [References](#) | [PDF \(3434 KB\)](#)

87
Views

0
CrossRef citations

0
Altmetric

CLINICAL PSYCHOLOGY & NEUROPSYCHOLOGY | REVIEW ARTICLE

[Parental marital quality and adolescent psychological well-being: A meta-analysis](#) >

Hepi Wahyuningsih, Fitri Ayu Kusumaningrum & Resnia Novitasari | Maria Cristina Gugliandolo (Reviewing editor)

Article: 1819005

Published online: 09 Sep 2020

[Abstract](#) | [Full Text](#) | [References](#) | [PDF \(472 KB\)](#)

275
Views

0
CrossRef citations

0
Altmetric



Psychometric assessment of an instrument evaluating the effects of affective variables on students' WTC in face-to-face and digital environment

Herri Mulyono , Regitha Saskia , Vista Sulaim Arrummaiza & Gunawan Suryoputro |

To cite this article: Herri Mulyono , Regitha Saskia , Vista Sulaim Arrummaiza & Gunawan Suryoputro | (2020) Psychometric assessment of an instrument evaluating the effects of affective variables on students' WTC in face-to-face and digital environment, Cogent Psychology, 7:1, 1823617

To link to this article: <https://doi.org/10.1080/23311908.2020.1823617>



© 2020 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.



Published online: 23 Sep 2020.



Submit your article to this journal [↗](#)



Article views: 146



View related articles [↗](#)



View Crossmark data [↗](#)



Received: 19 July 2020
Accepted: 07 September 2020

*Corresponding author: Herri Mulyono, University of Muhammadiyah Prof. DR. HAMKA, Jakarta Timur, Indonesia
E-mail: hmulyono@uhamka.ac.id

Reviewing editor:
Lucia Monacis, Università Degli Studi Di Foggia, Italy

Additional information is available at the end of the article

APPLIED PSYCHOLOGY | RESEARCH ARTICLE

Psychometric assessment of an instrument evaluating the effects of affective variables on students' WTC in face-to-face and digital environment

Herri Mulyono^{1*}, Regitha Saskia¹, Vista Sulaim Arrummaiza¹ and Gunawan Suryoputro¹

Abstract: The current study was aimed to translate and examine the psychometric characteristics of the Indonesian version of Lee and Hsieh's (2019) questionnaire measuring the effects of four affective variables (i.e. self-confidence, L2 anxiety, grit, and motivation) on students' willingness to communicate in face-to-face (F2F) in both inside and outside classrooms, and in a digital environment. Data were collected from 458 students—269 university students, 102 upper secondary school students, and 87 lower secondary school students. Several statistical analyses (i.e. unidimensionality assessment, reliability analysis, misfit analysis, and differential item functioning analysis) was performed using Rasch analysis. Findings of the study revealed that seven of eight constructs in the Indonesian version of Lee and Hsieh's (2019) questionnaire is valid with Indonesian samples, and they preserve the psychometric characteristics of the original scale. Recommendations are offered in reference to the findings.

ABOUT THE AUTHORS

Gunawan Suryoputro is a professor in English Education at the Graduate School, University of Muhammadiyah Prof. Dr. Hamka, Jakarta, Indonesia. Gunawan obtained her MA and PhD in applied linguistics from Atma Jaya Catholic University in Jakarta, Indonesia.

Herri Mulyono is a member of teaching staffs at English Department, Faculty of Teacher Training and Pedagogy, University of Muhammadiyah Prof. Dr. Hamka, Jakarta, Indonesia. Herri received his PhD in 2018 from University of York, UK. Herri currently chairs the L8 Education Research Group in the university.

Regitha Saskia graduated from the English Department, Faculty of Teacher Training and Pedagogy, University of Muhammadiyah Prof. Dr. Hamka, Jakarta, Indonesia. Currently, Regitha works as a research assistant in L8 Education Research Group in the university.

Vista Sulaim Arrummaiza is a student at the English Department, Faculty of Teacher Training and Pedagogy, University of Muhammadiyah Prof. Dr. Hamka, Jakarta, Indonesia. Currently, Vista works as a research assistant in L8 Education Research Group in the university.

PUBLIC INTEREST STATEMENT

A body of literature has suggested several constructs to examine students' willingness to communicate (WTC) in a foreign language classroom. However, little has been explored regarding the assessment of WTC constructs in face-to-face and digital settings, except one offered by Lee and Hsieh (2019)' questionnaire measuring the effects of four affective variables (i.e. self-confidence, L2 anxiety, grit, and motivation) on students' willingness to communicate (WTC) in face-to-face (F2F) settings (i.e. inside and outside classrooms) and in a digital environment. This current study attempted to examine the psychometric characteristics of the Indonesian version of Lee and Hsieh (2019) questionnaire using a Rasch analysis. The current study is significant in depicting an empirical evidence related to the psychometric characteristics of the Indonesian version of the Lee and Hsieh (2019) questionnaire using Rasch analysis. More importantly, it contributes to the under-explored topic in the literature on WTC in F2F and digital environment study within an Indonesian context.

Subjects: Educational Research; Language & Linguistics; Language Teaching & Learning;

Keywords: Rasch analysis; self-confidence; L2 anxiety; grit; motivation; willingness to communicate

1. Introduction

The term willingness to communicate (henceforth WTC) was first coined by McCroskey and Bear in 1985 (Khany & Nejad, 2017; Ningsih et al., 2018). The term was originally conceptualized to portray individual differences that encourage people to communicate in their first language (Yashima et al., 2018). In the second language (L2) learning context, the term L2 WTC is used to reflect one's readiness to take part in particular communication events using the target language when they are given opportunity (Khany & Nejad, 2017). Teachers' understanding of their students' WTC is crucial in L2 learning because students' WTC determine their participation in learning and achievement (Amiryousefi, 2018; Ningsih et al., 2018).

Many scholars have attempted to develop constructs to understand the nature of WTC. Horwitz et al. (1986) seminal work presents the Foreign Language Classroom Anxiety Scale (FLCAS), suggesting the connection of communication apprehension and students' learning anxiety in the classroom. Yashima (2002) synthesizes a number of literature on WTC constructs and reveals that WTC can predict students' communication frequency in L2, while students' motives can be used to predict their WTC, L2 communication frequency or both. MacIntyre, Clément et al. (1998) as cited in Yashima (2002) suggest several variables that influence students' L2 WTC or students' use of L2, including student personality, intergroup climate, attitudes, motivation, self-confidence, and communicative competence. Other latent variables such as students' cultural background, shyness, interaction issues are also known as factors that contribute to WTC (Cao, 2010) in addition to interest, motives and demographic factors (e.g. age and gender) (Amiryousefi, 2018) and potentially students' cognitive styles (de Sinatra et al., 2012).

Scholars have also examined the role of technology in promoting student WTC and students WTC in a digital environment. Some previous studies (e.g. Buckingham & Alpaslan, 2017; Lee, 2019; Reinders & Wattana, 2014; 2015; Waldeck et al., 2001) have revealed that the incorporation of technology in L2 learning had improved students' learning motivation and lower their affective obstructions; and thus increase student WTC. Several WTC constructs used in a few earlier studies, unfortunately still adopted ones developed for the assessment of WTC in nondigital setting, such as MacIntyre and Conrod (2001) and Cao and Philp (2006). Lee and associates (i.e. Lee, 2019; Lee & Hsieh, 2019; Lee & Lee, 2019) proposed a WTC construct that incorporated several affective variables (e.g. motivation, self-confidence, risk-taking, speaking anxiety, and grit), intercultural background, and WTC communication situation (inside the classroom, outside classroom, and in digital settings). Particularly, Lee and Hsieh (2019) develop a questionnaire to measure the effects of four affective variables (i.e. self-confidence, L2 anxiety, grit, and motivation) on students' willingness to communicate (WTC) in Face-to-face (F2F) settings (i.e. inside and outside classrooms) and in a digital environment.

This current study attempted to examine the psychometric characteristics of the Indonesian version of Lee and Hsieh (2019) questionnaire. Specifically, the current study was aimed to examine the internal construct of validity as well as the internal consistency of the translated questionnaire using the Rasch analysis. Rasch analysis is a psychometric technique used widely by instrument developers as well as researchers to monitor the quality of the instrument by estimating item difficulty and person ability at the same time (Yu, 2020). Rasch is believed to offer an objective measurement due to its capability in providing invariant measurement characteristics across diverse situations (Wright, 1992 as cited in Yu, 2020). To the best our knowledge, a validation study of Lee and Hsieh (2019) questionnaire using the Rasch analysis has not been reported elsewhere. The current study thus is significant in providing empirical evidence related to the psychometric characteristics of the Indonesian version of the questionnaire using Rasch

analysis. More importantly, it contributes to the underexplored topic in the literature on WTC in F2F and digital environment study within an Indonesian context.

2. Methods

2.1. Instrument

Lee and Hsieh (2019) questionnaire was aimed to measure the effects of four affective variables (i.e. self-confidence, L2 anxiety, grit, and motivation) on students' WTC in F2F in both inside and outside classrooms, and in a digital environment. The 33 questionnaire items were developed using a 5-point Likert scale and comprised of three sections: affective variable scale, second language willingness to communicate (L2 WTC) scale, and the demographic information of the participants. The affective variable subscale included six items of self-confidence subscale, six items of anxiety subscale, four items of motivation, and five items of grits. The other L2 WTC scale comprised of four items of F2F WTC inside classroom, four items of F2F WTC outside classroom, and four items of WTC in a digital environment. Responses to the questionnaire items in each subscale are detailed below.

2.2. Translation into Bahasa Indonesia

The original Lee and Hsieh (2019) questionnaire was written in English and translated into Bahasa Indonesia by the first and second author. The translated survey was sent to other researchers who were fluent both in English and Bahasa Indonesia, to validate and proofread the translation. Wordings of the Indonesian translation were refined in accordance to their feedback; ensuring the meaning of the original items was maintained in the translation.

2.3. Sample

Sample of the current study was selected using the following procedure: first, the researchers identified potential samples over social media (e.g. WhatsApp individual account or groups). Second, the invitations were sent to the target participants through a Google form link. Before completing the survey, the participants were asked to fill out a consent and demographic information session. A total of 458 students from lower secondary school, upper secondary school and Indonesian university completed the survey.¹ There were 337 females (73.6%) and 121 males (26.4%). Many of the participants were Indonesian EFL university students ($N = 269$, 58.7%), upper secondary school students ($N = 102$, 22.3%), and lower secondary school students ($N = 87$, 19%). Linacre (1994) argue that a sample size of 108 is considered appropriate in the case that the scale

Table 1. Response types for the subscale questionnaire items

Variables	Total item	Types of responses to the items
Self-confidence	6	Strongly disagree (1) to Strongly agree (5)
Speaking anxiety	6	Strongly disagree (1) to Strongly agree (5)
Motivation*	4	Strongly disagree (1) to Strongly agree (5)
Grit*	5	Not like me at all (1) to Very much like me (5)
F2F WTC inside classroom	4	Definitely not willing (1) to Definitely willing (5)
F2F WTC outside classroom	4	Definitely not willing (1) to Definitely willing (5)
WTC in a digital environment	4	Definitely not willing (1) to Definitely willing (5)

Note: One item in Motivation subscale and three items in Grit subscale were reverse coded.

is well-targeted, and a size of 243 is required for unspecific target. Given that the sample of 458 participants from three cohorts of Indonesian students was available in the current study, it is expected that the Rasch analysis can provide an appropriate level of precision from the data calculation. More importantly, the Outfit before Infit statistical analysis was performed to the data set to address the potential outliers that might violate the data and the analysis.

2.4. Analysis

Rasch analysis using WINSTEP (version 4.5.1) was performed to evaluate the data from 33 items. The analysis included several the assessment of an individual item and person fit through Outfit before Infit statistics and through mean square before Z-standardised (Zstd) fit (Ling Lee et al., 2020), unidimensionality, item and person separation reliability, the effectiveness of item scale, item and person mapping, and item bias (Chan & Subramaniam, 2020; Huang et al., 2020). The acceptable values of Infit and Outfit statistics range between 0.60 and 1.40; and between -2 and $+2$ for Zstd (Huang et al., 2020). A total of 169 samples were observed to be misfit and thus, were removed (Linacre, 2010). The remaining 289 samples included 169 university students (37%), 63 upper secondary school students (14%), and 58 lower secondary school students (12%). The total sample of 289 still met the recommended threshold for sample size (i.e. 50–250 samples) (Linacre, 1994).

3. Findings

3.1. The analysis of unidimensionality of the items

Unidimensionality assessment was carried out to evaluate if the questionnaire items measure a single construct (Yu, 2020). Unidimensionality assessment was done by evaluating the Rasch Principle Component Analysis (PCA) for each variable, and the finding revealed that 34.4%–64.9% of the variance is explained by the Rasch measures (see Table 2). The PCA range values were found greater than the PCA threshold of 20% (Sumintono & Widhiarso, 2014). Furthermore, the analysis of PCA showed ranges of eigenvalues of the first contrast for all variables were between 1.5 and 1.9; lower than 2.0 (Linacre, 2018). The findings have indicated that the Indonesian version of Lee and Hsieh (2019) questionnaire fitted to the Rasch model, reflecting a unidimensional measurement of the underlying construct.

3.2. Item and person separation reliability

Item separation reliability refers to the item reproducibility when the same items were given to a new sample with comparable ability, whereas person separation reliability concerns with the person classification reproducibility in a new sample when they respond the same items (Chang et al., 2014). The threshold of an item separation index was higher than 3 and with reliability higher than 0.90, while the person separation index should be greater than 2 with reliability greater than 0.8 (Linacre, 2018; Van Zile-Tamsen, 2017). As shown in Table 2, the item separation reliability of Lee and Hsieh (2019) questionnaire was excellent for the global scale and the subscales ($\alpha > 0.91$) and with high item separation for the global scale and the subscales (separation index > 3). The person separation reliability was considered good for the global scale ($\alpha = 0.80$), the subscale “Self-Confidence” ($\alpha = 0.80$), subscale “Speaking Anxiety” ($\alpha = 0.83$), scale “L2 WTC outside classroom” ($\alpha = 0.81$); and was observed at “fair to good” level for subscale “Motivation” ($\alpha = 0.62$), “L2 WTC inside classroom” ($\alpha = 0.78$), and “L2 WTC in digital environment” ($\alpha = 0.64$) (Bond & Fox, 2015; Ningrum et al., 2019). The person separation reliability of subscale “Grit” was very low ($\alpha = 0.44$). Although the person separation reliability of subscale “Motivation” ($\alpha = 0.62$), “L2 WTC inside classroom” ($\alpha = 0.78$), and “L2 WTC in digital environment” ($\alpha = 0.64$) were still considered fair to good internal consistency, the subscales are unlikely able to distinguish between high and lower performers from the relevant person sample (Linacre, 2018).

3.3. Effectiveness of the rating scales

The ordering functioning of the rating scale step categories was evaluated to understand participants ability to distinguish among the rating scales used in the questionnaire (see Table 1 for the rating scales). Table 3 below details the ordering functioning of the rating scale step categories

Parameter (with quality criteria)	Global-scale (33 items)	Self-	WTC outside classroom (WTC-Out)	WTC digital context (WTC-D)	confidence (SC)	Speaking anxiety (SA)
Motivation (M)	Grit (G)	WTC inside classroom (WTC-IC)				
<i>Model fit: summary of items</i>						
Item mean in logits (criteria: 0.0 logits)	.00, SD = .67	.00, SD = .85	.00, SD = .44	.00, SD = .58	.00, SD = .91	.00, SD = .47
Item reliability	.99	.98	.96	.96	.99	.93
Item separation reliability (criteria: good, 0.81–0.90; very good, 0.91–0.94; excellent, >0.94)	.99	.98	.96	.96	.99	.93
Item model fit MNSQ range extremes (criteria: good, 0.5–1.5; very good, 0.71–1.4; excellent, 0.77–1.3)	Infit: .54–1.67 Outfit: .56–1.68	Infit: .73–1.18 Outfit: .70–1.21	Infit: .80–1.40 Outfit: .77–1.39	Infit: .69–1.34 Outfit: .73–1.27	Infit: .91–1.15 Outfit: .89–1.17	Infit: .78–1.13 Outfit: .76–1.13
Item separation index (criteria: > 3)	8.89	6.98	4.97	4.72	12.05	3.70
Separate item strata = [(4 x separation index) + 1]/3 (criteria: fair, 2–3; good, 3–4; very good, 4–5; excellent, >5)	12.19 ≈ 12 levels	9.64 ≈ 10 levels	6.96 ≈ 7 levels	6.63 ≈ 7 levels	16.4 ≈ 16 levels	5.27 ≈ 5 levels
<i>Model fit: summary of persons</i>						
Person mean logits	.94, SD = .44	2.71, SD = 1.93	.59, SD = 1.55	3.51, SD = 1.55	.07, SD = .55	3.62, SD = 2.48
Person reliability	.80	.80	.83	.62	.44	.78
					3.43, SD = 2.66	2.93, SD = 1.72
					.81	.64

(Continued)

Parameter (with quality criteria)	Global-scale (33 items)	Self-	WTC outside classroom (WTC-Out)		WTC digital context (WTC-D)			confidence (SC)	Speaking anxiety (SA)
Motivation (M)	Grit (G)	WTC inside classroom (WTC-IC)							
Person separation reliability (criteria: good, 0.81–0.90; very good, 0.91–0.94; excellent, >0.94)	.80	.80	.83	.62	.44	.78	.81	.64	
Person separation index (criteria: > 2)	1.99	1.99	2.19	1.28	.88	1.91	2.09	1.35	
Separate item strata = [(4 x separation index) + 1]/3 (criteria: fair, 2–3; good, 3–4; very good, 4–5; excellent, >5)	2.99 \approx 3 levels	2.99 \approx 3 levels	3.25 \approx 3 levels	2.04 \approx 2 levels	1.51 \approx 2 levels	2.88 \approx 3 levels	3.12 \approx 3 levels	2.13 \approx 3 levels	
Rating scale analysis									
Responses per category (criteria: \geq 10)	Yes	NA	NA	NA	NA	NA	NA	NA	
Adjacent threshold distance (criteria: 1.4–5 logits)	(i) scale 0–1 was 0 to							–1.70 = 1.70; (ii) scale 1–2 was –1.70 (–.36) = 1.34 logits; (iii) scale 2–3 was –.36 –.40 = .76 logits; (iv) scale 3–4 was .40–1.67 = 1.27 logits	NA
NA	NA	NA	NA	NA	NA	NA	NA		
Outfit MNSQ (criteria: < 2 logits)	Yes	NA	NA	NA	NA	NA	NA	NA	
Probability curve graph (criteria: distinct curve on each response category and each peak is higher than 0.5 logit)	No	NA	NA	NA	NA	NA	NA	NA	

(Continued)

Table 2. (Continued)										
Parameter (with quality criteria)	Global-scale (33 items)	Self-	WTC outside classroom (WTC-Out)		WTC digital context (WTC-D)			confidence (SC)	Speaking anxiety (SA)	
Motivation (M)	Grit (G)	WTC inside classroom (WTC-IC)								
Average measure (criteria: increases monotonically across rating scale)	Yes	NA	NA	NA	NA	NA	NA	NA		
Dimensionality										
Raw variance in data explained by measures (criteria: > 20%)	34.4%	53.4%	61.6%	48.8%	50.6%	60.2%	64.9%	53.2%		
PCA eigenvalue for first contrast (criteria: > 2.0 indicates presence of another dimension; ≤ 2 supports unidimensional scale)	6.1	1.6	1.5	1.6	1.9	1.5	1.5	1.9		
Unexplained variance in 1st-5th contrasts of PCA of residuals (criteria: good, 5-10%; very good, 3-5%; excellent, < 3%)	3.3-12.1%	6.9-12.5%	6.0-9.3%	0.0-20.0%	0.1-19.1%	0.0-14.5%	0.0-13.5%	0.0-21.8%		

SD: standard deviation; PCA: principal components analysis; DIF: differential item functioning; MNSQ: mean square; NA: not applicable.

Table 3. Summary of category structure of 5-point rating scales

Rating scale step number	Observed person	Average calibration	Outfit MNSQ	Threshold	Threshold distance
1	231 (2)	-0.27	1.07	NONE	(-2.97)
2	1075 (11)	0.13	1.02	-1.7	-1.21
3	2212 (23)	0.62	0.92	-0.36	0.01
4	3545 (37)	1.02	1.19	0.4	1.22
5	2474 (26)	1.59	0.95	1.67	(-2.95)

As shown in the above Table 3, each of the rating scales included more than the minimum of 10 observations, and the distribution of responses by the category tended to be right-skewed. The outfit MNSQ values were observed less than 2.0, but the threshold distance between the rating scales was lower than the ideal threshold of 1.4–5 logits. Besides, the average calibration, as well as the step threshold, were observed to increase monotonically from the lowest rating point (i.e. -0.27) to the highest (1.59). These findings indicate that the category allowed the assessment of the latent variable and might suggest precision of the assessment (Van Zile-Tamsen, 2017). The monotonical increase in the average calibration suggested that “the higher category selected, the higher the students’ average measures” (DiStefano & Jiang, 2020, p. 39). Though, it was also found in the analysis that participants might have difficulty in distinguishing amongst the categories when completing the questionnaire.

3.4. Item and person mapping

Item and person maps (also known as Wright Maps) were developed to figure out the spread of subject perceptions and the distribution difficulty levels in the questionnaire items. Figure 1 presents the Wright item person map detailing the distribution of item and person difficulties. The left side of the map shows the distribution of the measured ability of the respondents from most able at the top to least able at the bottom, and the right side of the map shows the distribution of most difficult items at the top to the least difficult at the bottom. The item difficulties range from -1.21 logits to 1.61 logits.

As shown in Figure 1, item Q18 (-1.21 logit) of subscale “Grit” was observed to be the hardest item to agree amongst the students, whereas item Q16 (1.61 logit) of subscale “Motivation” was the easiest item to agree. The subscale “Grit” had three difficulty items to answer by the participants (i.e. Q18, Q20, Q21) that they were two standard deviations above the mean item difficulty. Three items in the subscale “Motivation” (i.e. Q14, Q 15, and Q16) and subscale WTC in a digital environment (i.e. Q30, Q31 and Q32) were easy to answer as they were two standard deviations below the mean item difficulty level. The findings indicate that the assessment instrument is likely easy to measure students’ motivation to communicate in L2, but is relatively difficult to identify the grit factor of the group of students.

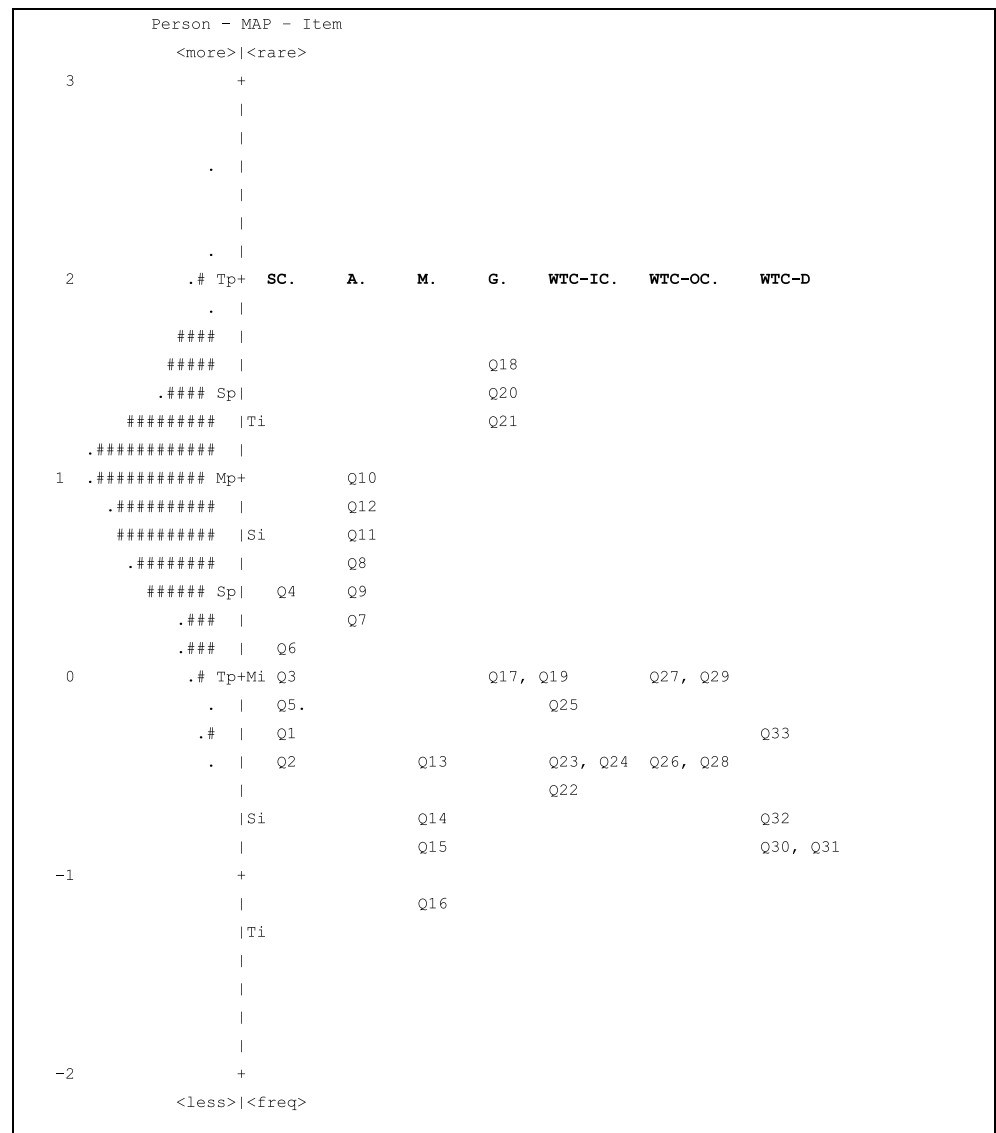
3.5. Item bias

The analysis of differential item functioning (DIF) using Rasch-Welch tests was performed to evaluate if there were potential item biases caused by group characteristics (e.g. gender and level of education) (Chan & Subramaniam, 2020). An item is considered as DIF if the value of DIF contrast (the difference in difficulty of the questionnaire item) is greater than 0.5 logits and were significant (Rasch-Welch probability value < 0.05) (Chan & Subramaniam, 2020; Linacre, 2018). DIF evaluation of the participant gender revealed all DIF contrast values were lower than 0.5 logits. In contrast to gender characteristic, several items with significant DIF were found in reference to students’ level of education (see Table 4).

Item Q2 was considered too easy for lower secondary school students, but not for university students (DIF LS = -0.86 logit, DIF U = -0.31 logit, $p < 0.05$). Item Q4 was observed to ease lower secondary school students, but was considered difficult for upper secondary and university

Figure 1. Wright person-item map (N = 289).

“#” represents three persons;
 “.” Represents one to two persons. M_p : person mean; S_p : one standard deviation of person mean; T_p : two standard deviation of person mean; M_i : item mean; S_i : one standard deviation of item mean; T_i : two standard deviation of item mean; self-confidence (SC) items: Q1-6; Speaking anxiety (A): Q7-12; Motivation (M): Q13-16; Grit (G): Q17-21; WTC inside classroom (WTC-IC): Q22-25; WTC outside classroom (WTC-OC): Q26-29; WTC in digital environment (WTC-D): Q30-33.



students (DIF LS = -0.04 logit, DIF US = 0.49 logit, DIF U = 0.55 logit, $p < 0.05$). University students perceived easy to complete item Q9 while lower secondary school students considered item Q18 as an easy item (DIF U = 0.22 logit, DIF LS = 1.24 logit, $p < 0.05$)

4. Discussion and conclusion

Although a quite number of current literatures on WTC offers measures to evaluate students' willingness to communicate in L2 setting, they rarely compared two different WTC settings: in F2F classroom settings and in digital environment. The current study examined the psychometric characteristics of the Indonesian version of Lee and Hsieh (2019) questionnaire measuring the effects of four affective variables (i.e. self-confidence, L2 anxiety, grit, and motivation) on students' willingness to communicate (WTC) in both F2F inside and outside classrooms and in a digital environment. It contributes to the under-explored topic in the literature on WTC in F2F and digital environment study within an Indonesian context.

Table 4. DIF analysis by the level of education

Item	Level of education	DIF measure	DIF Contrast	t	Probability
Q2	U	−0.31	0.54	2.70	0.00
	LS	−0.86			
Q4	US	0.49	0.53	2.61	0.01
	LS	−0.04			
	U	0.55	0.59	3.41	0.00
	LS	−0.04			
Q9	LS	0.86	0.63	0.17	0.00
	U	0.22			
Q18	U	1.75	0.52	3.04	0.00
	LS	1.24			

Note. LS = lower secondary school, US = upper secondary school, U = university

The psychometric properties of the Indonesian version of Lee and Hsieh (2019) questionnaire were examined through several different types of analysis (i.e. analyses of unidimensionality of the items, item and person separation reliability, item and person mapping, and item bias), primarily based upon item response theory (IRT). IRT theory is concerned with the assessment of individual items, offering person and item parameter invariance when the model fit is presented, and providing more detailed information about the scale measuring particular constructs of interest (Zanon et al., 2016). Based upon such a theory, Rasch analysis provides researchers with a tool to assess the psychometric quality of the scale and enable them to improve particular aspect, allowing them to make a prediction of person scores according to his/her item performance (Boone et al., 2014; Linacre, 2018; Zanon et al., 2016).

The result from Rasch analysis has shown that Lee and Hsieh (2019) questionnaire possess a sufficient range of item difficulty. The monotonical increase found on the average calibration of the rating points has indicated that the rating scale used in the questionnaire had allowed the assessment of the eight variables and might suggest precision of the assessment. Of the eight variables in Lee and Hsieh (2019) questionnaire, six were considered to have strong internal consistency and one with a fair internal consistency level. The other variable “grit” had a low-reliability level and thus, should not be used in the questionnaire. Findings of the current also suggest that three items in grit subscale were of the most difficult items to agree. These findings corresponded our early analysis (see Mulyono & Saskia, 2020), showing that the variable “motivation” was marginally reliable, and the variable “grit” subscale had unacceptably low reliability.

Although a body of literature has provided strong empirical evidence about the role of grit in promoting students’ WTC in a foreign language (Lee, 2020; Lee & Draji, 2019; Lee & Lee, 2019), Credé et al. (2017) have raised a concern regarding the construct validity of scales measuring the individual level of grit. In their meta-analysis on the study about grit, Credé et al. (2017) have identified several empirical evidence suggesting that the relationship between individual’s grit and his/her performance might not be strong in any types of settings. By citing MacNamara et al. (2014), Credé, Tynan, and Harms argue that the high level of individual’s grit may only be obtained when he/she is given difficult but clearly defined tasks. Such difficult but clear tasks are unlikely well articulated in the construct of grit offered by Lee and Hsieh (2019). It is thus, items in grit needs to be redesigned to reflect “perseverance of effort” construct (Credé et al., 2017; Lee, 2020).

Overall, the Indonesian version of Lee and Hsieh (2019) questionnaire measuring the effects of four affective variables (i.e. self-confidence, L2 anxiety, grit, and motivation) on students’ WTC in F2F settings (i.e. inside and outside classrooms) and in a digital environment fit the proposed

model based on Rasch analysis. The questionnaire is substantiated by no item bias found across male and female participants. However, four items with DIF were found across the background of the participants (i.e. Q2, Q4, Q9, and Q18), particularly between university students and lower secondary school students. Rephrasing these four items is recommended to address potential item bias for participants with different educational background. More description of the context of F2F outside the classroom and digital settings is also suggested in the future update of the questionnaire.

Acknowledgments

The authors would like to thank Dr. Bambang Sumintono for his thoughts and feedback on the Rasch analysis result.

Funding

The publication of this article is funded by the Scientific Publication Advancement and Support Unit, University of Muhammadiyah Prof. DR. HAMKA, Jakarta, Indonesia under Publication Incentive Scheme.

Author details

Herri Mulyono¹
 E-mail: E-mail: hmulyono@uhamka.ac.id
 ORCID ID: <http://orcid.org/0000-0002-0668-0268>
 Regitha Saskia¹
 Vista Sulaim Arrummaiza¹
 Gunawan Suryoputro¹
 ORCID ID: <http://orcid.org/0000-0001-6841-5753>

¹ University of Muhammadiyah Prof. DR. HAMKA, Jakarta Timur, Indonesia.

Conflicts of interest

The authors declare no conflict of interest.

Citation information

Cite this article as: Psychometric assessment of an instrument evaluating the effects of affective variables on students' WTC in face-to-face and digital environment, Herri Mulyono, Regitha Saskia, Vista Sulaim Arrummaiza & Gunawan Suryoputro, *Cogent Psychology* (2020), 7: 1823617.

Note

1. The data are published and available in Mulyono and Saskia (2020).

References

- Amiryousefi, M. (2018). Willingness to communicate, interest, motives to communicate with the instructor, and L2 speaking: A focus on the role of age and gender. *Innovation in Language Learning and Teaching*, 12(3), 221–234. <https://doi.org/10.1080/17501229.2016.1170838>
- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). Routledge.
- Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch analysis in the human sciences*. Springer.
- Buckingham, L., & Alpaslan, R. S. (2017). Promoting speaking proficiency and willingness to communicate in Turkish young learners of English through asynchronous computer-mediated practice. *System*, 65, 25–37. <https://doi.org/10.1016/j.system.2016.12.016>
- Cao, Y. (2010). Investigating situational willingness to communicate within second language classrooms from an ecological perspective. *System*, 39(4), 468–479. <https://doi.org/10.1016/j.system.2011.10.016>
- Cao, Y., & Philp, J. (2006). Interactional context and willingness to communicate: A comparison of behavior in whole class, group and dyadic interaction. *System*, 34(4), 480–493. <https://doi.org/10.1016/j.system.2006.05.002>
- Chan, M., & Subramaniam, R. (2020). Validation of a science concept inventory by Rasch Analysis. In M. Khine (Ed.), *Rasch measurement: Applications in quantitative educational research* (pp. 159–178). Springer.
- Chang, K.-C., Wang, J.-D., Tang, H.-P., Cheng, C.-M., & Lin, C.-Y. (2014). Psychometric evaluation, using Rasch analysis, of the WHOQOL-BREF in heroin-dependent people undergoing methadone maintenance treatment: Further item validation. *Health and Quality of Life Outcomes*, 12(1), 148. <https://doi.org/10.1186/s12955-014-0148-6>
- Credé, M., Tynan, M. C., & Harms, P. D. (2017). Much ado about grit: A meta-analytic synthesis of the grit literature. *Journal of Personality and Social Psychology*, 113(3), 492–511. <https://doi.org/10.1037/pspp0000102>
- de Sinatra, P. V., Tanucci, M., Monacis, G., Bitonto, L., Di, P., Roselli, T., & Rossano, V. (2012). How cognitive styles affect the E-learning process. In 2012 IEEE 12th international conference on advanced learning technologies (pp. 359–363). Rome, Italy. <https://doi.org/10.1109/ICALT.2012.79>
- DiStefano, C., & Jiang, N. (2020). Applying the Rasch rating scale method to questionnaire data. In M. S. Khine (Ed.), *Rasch measurement: Applications in quantitative educational research* (pp. 31–46). Springer.
- Horwitz, E. K., Horwitz, M., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70(2), 125–132. <https://doi.org/10.1111/j.1540-4781.1986.tb05256.x>
- Huang, F., Huang, L., & Oon, P.-T. (2020). Constructs evaluation of student attitudes towards science - A Rasch analysis. In M. S. Khine (Ed.), *Rasch measurement: Applications in quantitative educational research* (pp. 139–158). Springer.
- Khany, R., & Nejad, A. M. (2017). L2 willingness to communicate, openness to experience, extraversion, and L2 unwillingness to communicate: The Iranian EFL context. *RELJ Journal*, 48(2), 241–255. <https://doi.org/10.1177/0033688216645416>
- Lee, J. S. (2019). EFL students' views of willingness to communicate in the extramural digital context. *Computer Assisted Language Learning*, 32(7), 692–712. <https://doi.org/10.1080/09588221.2018.1535509>
- Lee, J. S. (2020). The role of grit and classroom enjoyment in EFL learners' willingness to communicate. *Journal of Multilingual and Multicultural Development*, 1–17. <https://doi.org/10.1080/01434632.2020.1746319>
- Lee, J. S., & Drajiati, N. A. (2019). Affective variables and informal digital learning of English: Keys to willingness to communicate in a second language. *Australasian Journal of Educational Technology*, 35(5), 168–182. <https://doi.org/10.14742/ajet.5177>
- Lee, J. S., & Hsieh, J. C. (2019). Affective variables and willingness to communicate of EFL learners in in-class, out-of-class, and digital contexts. *System*, 82, 63–73. <https://doi.org/10.1016/j.system.2019.03.002>

- Lee, J. S., & Lee, K. (2019). Affective factors, virtual inter-cultural experiences, and L2 willingness to communicate in in-class, out-of-class, and digital settings. *Language Teaching Research*, 24(6), 1–21. <https://doi.org/10.1177/1362168819831408>
- Linacre, J. M. (1994). Sample size and item calibration stability. *Rasch Measurement Transactions*, 7(4), 328. <https://www.rasch.org/rmt/rmt74m.htm>
- Linacre, J. M. (2010). When to stop removing items and persons in Rasch misfit analysis? *Rasch Measurement Transactions*, 23(4), 1241. <https://www.rasch.org/rmt/rmt234g.htm>
- Linacre, J. M. (2018). A user's guide to Winsteps Ministep Rasch-model computer programs (version 4.3.1). Institute for Objective Measurement, Inc. Retrieved March 19, 2019, from <https://www.winsteps.com/a/Winsteps-Manual.pdf>
- Ling Lee, W., Chinna, K., & Sumintono, B. (2020). Psychometrics assessment of HeartQoL questionnaire: A Rasch analysis. *European Journal of Preventive Cardiology*, 204748732090232. <https://doi.org/10.1177/2047487320902322>
- MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal*, 82(4), 545–562
- MacIntyre, P. D., & Conrad, S. (2001). Willingness to communicate, social support, and language-learning orientations of immersion students. *Studies in Second Language Acquisition*, 23(3), 369–388. <https://doi.org/10.1017/S0272263101003035>
- Macnamara, B. N., Hambrick, D. Z., & Oswald, F. L. (2014). Deliberate Practice and Performance in Music, Games, Sports, Education, and Professions: A Meta-Analysis. *Psychological science*, 25(8), 1608–1618. doi:10.1177/0956797614535810
- Mulyono, H., & Saskia, R. (2020). Dataset on the effects of self-confidence, motivation and anxiety on Indonesian students' willingness to communicate in face-to-face and digital settings. *Data in Brief*, 31, 105774. <https://doi.org/10.1016/j.dib.2020.105774>
- Ningrum, E., Evans, S., & Soh, S.-E. (2019). Validation of the Indonesian version of the Safety Attitudes Questionnaire: A Rasch analysis. *PloS One*, 14(4), e0215128. <https://doi.org/10.1371/journal.pone.0215128>
- Ningsih, S. K., Narahara, S., & Mulyono, H. (2018). An exploration of factors contributing to students' unwillingness to communicate in a foreign language across Indonesian secondary schools. *International Journal of Instruction*, 11(4), 811–824. <https://doi.org/10.12973/iji.2018.11451a>
- Reinders, H., & Wattana, S. (2014). Can I say something? The effects of digital game play on willingness to communicate. *Language Learning & Technology*, 18(2), 101–123. <https://hdl.handle.net/10652/2962>
- Reinders, H., & Wattana, S. (2015). Affect and willingness to communicate in digital game-based learning. *ReCALL*, 27(1), 38–57. <https://doi.org/10.1017/S0958344014000226>
- Sumintono, B., & Widhiarso, W. (2014). *Aplikasi Model Rasch untuk penelitian ilmu-ilmu sosial* (2nd ed.). Trim Komunikata.
- Van Zile-Tamsen, C. (2017). Using Rasch analysis to inform rating scale development. *Research in Higher Education*, 58(8), 922–933. <https://doi.org/10.1007/s11162-017-9448-0>
- Waldeck, J., Kearney, P., & Plax, T. (2001). Teacher e-mail message strategies and students' willingness to communicate online. *Journal of Applied Communication Research*, 29(1), 54–70. <https://doi.org/10.1080/00909880128099>
- Wright, B.D. (1992). The international objective measurement workshops: Past and future. In M. Wilson (Ed), *Objective measurement: Theory into practice* (Vol. 1, pp. 9–28). Ablex Publishing.
- Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *The Modern Language Journal*, 86(1), 54–66. <https://doi.org/10.1111/1540-4781.00136>
- Yashima, T., MacIntyre, P. D., & Ikeda, M. (2018). Situated willingness to communicate in an L2: Interplay of individual characteristics and context. *Language Teaching Research*, 22(1), 115–137. <https://doi.org/10.1177/1362168816657851>
- Yu, H. C. (2020). Objective measurement: How Rasch modeling can simply and enhance your assessment. In M. Khine (Ed.), *Rasch measurement: Applications in quantitative educational research* (pp. 47–73). Springer.
- Zanon, C., Hutz, C. S., Yoo, H. (Henry), & Hambleton, R. K. (2016). An application of item response theory to psychological test development. *Psicologia: Reflexão E Crítica*, 29(1), 18. <https://doi.org/10.1186/s41155-016-0040-x>



© 2020 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

You are free to:

Share — copy and redistribute the material in any medium or format.

Adapt — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made.

You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions

You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Cogent Psychology (ISSN:) is published by Cogent OA, part of Taylor & Francis Group.

Publishing with Cogent OA ensures:

- Immediate, universal access to your article on publication
- High visibility and discoverability via the Cogent OA website as well as Taylor & Francis Online
- Download and citation statistics for your article
- Rapid online publication
- Input from, and dialog with, expert editors and editorial boards
- Retention of full copyright of your article
- Guaranteed legacy preservation of your article
- Discounts and waivers for authors in developing regions

Submit your manuscript to a Cogent OA journal at www.CogentOA.com

